

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 23:15:55 ; Search time 166 Seconds
(without alignments)
368.121 Million cell updates/sec

Title: US-09-525-041-2

Perfect score: 878
Sequence: 1 MASSRMILLILSLCAKTGV.....LTWSNECKRQHFICKTRP 158

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq16dec04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878	100.0	158	2	AAW12691 Human col
2	878	100.0	158	2	AAW37929 A human R
3	878	100.0	158	2	AAW37866 Human pro
4	878	100.0	158	2	AAW84274 Protein e
5	878	100.0	158	3	AAW12900 Human col
6	878	100.0	158	4	AAW74934 Human TSA
7	878	100.0	158	4	AAW24519 C880P sim
8	878	100.0	158	4	AAW24520 C880P sim
9	878	100.0	158	4	AAW24521 C880P sim
10	878	100.0	158	4	AAW24517 C81-152 C
11	878	100.0	158	5	AAW24518 C880P sim
12	878	100.0	158	5	AAW29829 Human REG
13	878	100.0	158	5	AAW29893 Human REG
14	878	100.0	158	6	ABR58851 Human can
15	878	100.0	158	6	ABP56022 Human REG
16	878	100.0	158	6	ADA83832 Human REG
17	878	100.0	158	6	ABP55366 Human col
18	878	100.0	158	6	ABP55370 Human col
19	878	100.0	158	6	ABP55369 Human col
20	878	100.0	158	6	ABP55367 Human col
21	878	100.0	158	6	ABP55368 Human col
22	878	100.0	158	7	ADB80529 Ovarian c
23	878	100.0	158	7	ADB75523 Prostate
24	878	100.0	158	7	ADF16659 Human alb
25	878	100.0	158	7	ADF56454 Human col

26	878	100.0	158	7	ADF85474 Human reg
27	878	100.0	158	7	ADH21872 Human Reg
28	878	100.0	158	7	ADN39463 Cancer/an
29	878	100.0	158	7	ADN39882 Cancer/an
30	878	100.0	158	7	ADN38820 Cancer/an
31	878	100.0	158	7	ADN39543 Cancer/an
32	878	100.0	158	8	AD158211 Human reg
33	878	100.0	158	8	ADN35904 Human reg
34	878	100.0	158	8	ADQ29679 Human col
35	878	100.0	158	8	ADQ80359 Regenerat
36	878	100.0	158	8	ADP55956 Human PRO
37	878	100.0	166	4	AAW75620 Human col
38	878	100.0	767	7	ADF16496 Human alb
39	878	100.0	767	7	ADH21795 Human alb
40	867	98.7	158	5	ABR99310 Amino aci
41	856	97.5	153	3	AAW92267 Human can
42	760	86.6	367	8	ADF73150 RELP-Fc F
43	695	79.2	122	6	ABP76307 Human GEN
44	307.5	35.0	134	5	ABP59097 Human zin
45	254.5	29.0	165	1	ABP81513 Sequence

ALIGNMENTS

RESULT 1
AAW12691 standard; protein; 158 AA.
XX
XX
AC AAW12691;
XX
XX
DT 31-MAY-1997 (first entry)
XX
XX
DE Human colon specific protein.
XX
XX
KW Colon specific protein; colon cancer; metastasis; diagnosis; therapy;
KW antibody; vaccine; agonist; antagonist.
XX
XX
OS Homo sapiens.
XX
XX
PN WO9639541-A1.
XX
XX
PD 12-DEC-1996.
XX
XX
PF 06-JUN-1995; 95WO-US007169.
XX
XX
PR 06-JUN-1995; 95WO-US007169.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Soppet DR, Li Y, Dillon PJ;
XX
XX
DR WPI-4397-043162/04.
XX
XX
PT N-PSDB; AAT51784.
XX
XX
PT New isolated colon specific gene - used to develop prods. for use in the
XX
XX
PT diagnosis and treatment of colon disorders, partic. colon cancer.
XX
XX
PS Claim 1; Page 53; 64pp; English.
XX
XX
CC A human colon specific protein (AAW12691) is a potential diagnostic
CC marker for colon cancer. It is believed that the presence of active
CC transcriptions of the colon specific gene in non-colon cells of a host is
CC indicative of colon cancer metastases. The amino acid sequence of the
CC colon specific protein was deduced from a cDNA clone (AA151784) isolated
CC from a human colon cancer cDNA library. Recombinant colon specific
CC protein can be produced in transformed host (e.g. bacterial, insect)
CC cells and used to develop prods. for the diagnosis and treatment of colon
CC disorders, partic. colon cancer metastasis. Antibodies raised against the
CC protein can be used to target cancer cells and as part of a colon cancer
CC vaccine
XX
SQ Sequence 158 AA;

Query Match 100.0%; Score 878; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLLSCIAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNMSDALEECQS 60
DB 1 MASRMRLLLLLSCIAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNMSDALEECQS 60
QY 61 YGNNGHLLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKROQOWIDGAMTYLRSWSG 120
DB 61 YGNNGHLLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKROQOWIDGAMTYLRSWSG 120
QY 121 KSMGNGKHCALMSNNNFLTWSNENCRORHFLCKYRP 158
DB 121 KSMGNGKHCALMSNNNFLTWSNENCRORHFLCKYRP 158

RESULT 2
AAM37929
ID AAM37929 standard; protein; 158 AA.

XX AAM37929;
XX
XX
XX 21-AUG-1998 (first entry)

DE A human Reg I-gamma protein.

XX Human; Reg I-gamma protein; C-type lectin; Incyte clone 1310334;
XX regulation; cell growth; development; tumorigenesis; neurodegeneration;
XX inhibition; treatment; prevention; neoplasia; metastasis;
XX neurodegenerative change; Alzheimer's disease; Down's syndrome;
XX regeneration; pancreatic beta-cells; diabetes.

XX Homo sapiens.
XX
XX MO9816640-A1.

XX
XX 23-APR-1998.

XX 07-OCT-1997; 97WO-US018174.

XX 11-OCT-1996; 96US-00729103.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Goli SK;

XX WPI, 1998-251287/22.

XX N-PSDB; AAV29156.

XX New isolated human Reg I-gamma protein useful for developing products
XX for treating, e.g. diabetes, tumours, or neuro-degenerative disease such
XX as Alzheimer's.

XX Claim 1; Fig 1; 72pp; English.

XX The present sequence represents a human Reg I-gamma protein, which
XX comprises a C-type lectin. The sequence was identified in Incyte clone
XX 1310334. Reg I-gamma protein is involved in regulation of cell growth and
XX development. Since the overexpression of reg protein is associated with
XX tumorigenesis and neurodegeneration, inhibition of human Reg I-gamma
XX expression can be used for treating or preventing neoplasia or metastasis
XX and neurodegenerative changes associated with Alzheimer's disease and
XX other disorders of the central nervous system, e.g. Down's syndrome. Reg
XX I-gamma can also be used in therapeutics to induce regeneration of
XX pancreatic beta-cells in the treatment of diabetes. The products can also
XX be used for detection for, e.g. expression of REG I-gamma, diagnosis and
XX drug screen

XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 2; Length 158;

Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLLSCIAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNMSDALEECQS 60
DB 1 MASRMRLLLLLSCIAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNMSDALEECQS 60
QY 61 YGNNGHLLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKROQOWIDGAMTYLRSWSG 120
DB 61 YGNNGHLLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKROQOWIDGAMTYLRSWSG 120
QY 121 KSMGNGKHCALMSNNNFLTWSNENCRORHFLCKYRP 158
DB 121 KSMGNGKHCALMSNNNFLTWSNENCRORHFLCKYRP 158

RESULT 3
AAM37866
ID AAM37866 standard; protein; 158 AA.

XX AAM37866;

XX 10-AUG-1998 (first entry)

DE Human protein comprising secretory signal amino acid sequence 3.

XX Human protein; secretory signal; nutritional source; cytokine; immunity;
XX haematopoiesis; activin; inhibin; tumour; chemotactic; chemokinetic;
XX thrombolytic; anti-inflammatory; inhibition; stomach cancer cell.

XX Homo sapiens.

XX WO9811217-A2.

XX 19-MAR-1998.

XX 12-SEP-1997; 97WO-JP003239.

XX 13-SEP-1996; 96JP-00243060.

XX (SAGA) SAGAMI CHEM RES CENTRE.

XX (PROT-) PROTEGENE INC.

XX Kato S, Sekine S, Yamaguchi T, Kobayashi M;

XX WPI, 1998-207380/18.

XX N-PSDB; AAV29035; AAV29036.

XX Human proteins with secretory signal sequences - used to treat immune
XX deficiencies, infections, tumours, and haematopoietic disorders, etc.

XX Claim 1; Page 67-68; 13pp; English.

XX This is the amino acid sequence of a novel human protein comprising a
XX secretory signal isolated from stomach cancer cells. Its proteins can be
XX used as nutritional sources or supplements. The proteins may also have
XX cytokine functions, immune modulating functions, haematopoiesis
XX regulating activity, activin/inhibin regulating activity,
XX chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
XX receptor/ligand activity, anti-inflammatory activity, tumour inhibition
XX activity

XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLLSCIAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNMSDALEECQS 60
DB 1 MASRMRLLLLLSCIAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNMSDALEECQS 60
QY 61 YGNNGHLLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKROQOWIDGAMTYLRSWSG 120

DB 61 YGNGAHLASTILSKASTAETAYISGYORSQPIWIGLHPQKQOMWIDGAMTYLRSWSG 120
QY 121 KSMGNKHCACEMSSNNNFLTWSNCCNKRQHFLCKTRP 158
DB 121 KSMGNKHCACEMSSNNNFLTWSNCCNKRQHFLCKTRP 158

RESULT 4

AAW84274
ID AAW84274 standard; protein; 158 AA.

AC AAW84274;

DT 25-MAR-1999 (first entry)

DE Protein encoded by a human colon specific gene.

KM Human; colon specific gene; diagnosis; colon disorder; colon cancer;
viability; colon cancer cell.

OS Homo sapiens.

PN US5861494-A.

PD 19-JAN-1999.

PF 06-JUN-1995; 95US-00468413.

PR 06-JUN-1995; 95US-00468413.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Dillon PJ, Soppet DR, Li Y;

DR WPI; 1999-130432/11.

DR N-PSDB; AAX03195.

PT Isolated human colon specific gene - used to develop products for the
diagnosis and treatment of disorders of the colon, e.g. colon cancer and
metastases.

PS Claim 1; Fig 1A-C; 20pp; English.

CC The present sequence is encoded by a human colon specific gene. The
nucleic acid sequence can be used to develop products for the diagnosis
of a disorder of the colon, e.g. colon cancer or metastases. The products
can also be used to screen for agonists or antagonists for the
polypeptides. The antagonists may be used to treat colon cancer, since
they interact with the function of colon specific polypeptides to inhibit
functions which are necessary for the viability of colon cancer cells.
CC The products can also be used for the production of antibodies and for
the identification of receptors for the polypeptides

XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;

Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLILSLCAKTGVLGDIIMRPSCAPGWFYHKSNCYGFPRKLRNMSDALEECOS 60
DB 1 MASRSMRLLILSLCAKTGVLGDIIMRPSCAPGWFYHKSNCYGFPRKLRNMSDALEECOS 60

QY 61 YGNGAHLASTILSKASTAETAYISGYORSQPIWIGLHPQKQOMWIDGAMTYLRSWSG 120
DB 61 YGNGAHLASTILSKASTAETAYISGYORSQPIWIGLHPQKQOMWIDGAMTYLRSWSG 120

QY 121 KSMGNKHCACEMSSNNNFLTWSNCCNKRQHFLCKTRP 158
DB 121 KSMGNKHCACEMSSNNNFLTWSNCCNKRQHFLCKTRP 158

RESULT 5
ID AAB12900
ID AAB12900 standard; protein; 158 AA.

AC AAB12900;

DT 14-NOV-2000 (first entry)

DE Human colon specific protein sequence.

KM Human; colon specific; colon cancer; metastasis; diagnosis; treatment;
cytostatic.

OS Homo sapiens.

PN US6080722-A.

PD 27-JUN-2000.

PF 29-SEP-1998; 98US-00162508.

PR 06-JUN-1995; 95US-00468413.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Dillon PJ, Soppet DR, Li Y;

DR WPI; 2000-464055/40.

DR N-PSDB; AAA62951.

PT Novel human colon specific polypeptides and polynucleotides for diagnosis
and treatment of colon cancer, for screening compounds which interact
with polypeptide, for synthesis of DNA and manufacture of DNA vectors.

PS Claim 1; Fig 1; 20pp; English.

CC This invention relates to a purified human protein, which is primarily
expressed in tissue derived from the colon. The protein is 152 amino
acids in length and exhibits cytostatic activity. The present sequence
represents the amino acid sequence of the colon specific protein. The
CC protein can be used in the diagnosis and treatment of colon cancer, and
it is thought that abnormally high levels of the gene expression in non-
colon cells is an indication of colon cancer metastasis

XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;

Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLILSLCAKTGVLGDIIMRPSCAPGWFYHKSNCYGFPRKLRNMSDALEECOS 60
DB 1 MASRSMRLLILSLCAKTGVLGDIIMRPSCAPGWFYHKSNCYGFPRKLRNMSDALEECOS 60

QY 61 YGNGAHLASTILSKASTAETAYISGYORSQPIWIGLHPQKQOMWIDGAMTYLRSWSG 120
DB 61 YGNGAHLASTILSKASTAETAYISGYORSQPIWIGLHPQKQOMWIDGAMTYLRSWSG 120

QY 121 KSMGNKHCACEMSSNNNFLTWSNCCNKRQHFLCKTRP 158
DB 121 KSMGNKHCACEMSSNNNFLTWSNCCNKRQHFLCKTRP 158

RESULT 6

ID AAB74934
ID AAB74934 standard; protein; 158 AA.

AC AAB74934;

DT 27-JUN-2001 (first entry)

DE Human TSA7005 protein SEQ ID NO:1.

KM Human; TSA7005; Reg; pancreatic beta cell growth; hypoglycaemic;
 KM diagnosis.
 XX
 OS Homo sapiens.
 XX JP2001025389-A.
 XX
 XX 30-JAN-2001.
 XX
 XX 15-JUL-1999; 99JP-00201279.
 XX
 XX 15-JUL-1999; 99JP-00201279.
 XX
 XX (SAKA) OTSUKA PHARM CO LTD.
 XX
 DR WPI: 2001-303742/32.
 DR N-ESDB; AAF82117, AAF82118.
 XX
 PT TSA7005 gene, encoding a polypeptide useful for the diagnosis and
 PT treatment of diseases associated with its expression.
 XX
 PS Claim 1; Page 23; 25pp; Japanese.
 XX
 CC The present sequence represents a human TSA7005 protein which shares 32%
 CC homology with human and mouse Reg proteins, and 34% homology with the rat
 CC Reg protein. TSA7005 has pancreatic beta cell growth activity and
 CC hypoglycaemic activity. The TSA7005 protein can be used for the diagnosis
 CC and treatment of diseases associated with the gene and its expression
 CC product
 XX
 SQ Sequence 158 AA;
 XX
 Query Match 100.0%; Score 878; DB 4; Length 158;
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASRSRLLLLSCLAKTGVLDIMRPSCAPGWFYHKSNCYGFRLKRNMSDALEEQS 60
 DB 1 MASRSRLLLLSCLAKTGVLDIMRPSCAPGWFYHKSNCYGFRLKRNMSDALEEQS 60
 QY 61 YGNGAHLASIIISLKEASTIAEYISGYORSQPIWIGLHPQKROQWIDGAMLYRWSWG 120
 DB 61 YGNGAHLASIIISLKEASTIAEYISGYORSQPIWIGLHPQKROQWIDGAMLYRWSWG 120
 QY 121 KSMGNGKHCAMSSNNNFLTWSNCKRQHFLLCKYRP 158
 DB 121 KSMGNGKHCAMSSNNNFLTWSNCKRQHFLLCKYRP 158
 RESULT 7
 AAM24519
 ID AAM24519 standard; protein; 158 AA.
 XX
 AC AAM24519;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE C880P similar amino acid sequence (GENESEQ W37866).
 XX
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
 KW gene therapy; vaccine; colonic cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PE 29-DEC-2000; 2000WO-US0355596.
 XX
 PR 30-DEC-1999; 99US-00476296.
 PR 10-JAN-2000; 2000US-00480321.
 PR 15-FEB-2000; 2000US-00504629.

PR 06-MAR-2000; 2000US-00519444.
 PR 19-MAY-2000; 2000US-00575251.
 PR 29-JUN-2000; 2000US-00609448.
 PR 28-AUG-2000; 2000US-00649811.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 PI King GR, Wang T, Jiang Y;
 XX
 DR WPI: 2001-441847/47.
 XX
 PT Colon tumor associated proteins and nucleic acids useful for the
 PT prevention, diagnosis and treatment of colonic cancer.
 XX
 PS Claim 2; Page 467-468; 472pp; English.
 XX
 CC The present invention describes colon tumour associated proteins (I) and
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate colon tumour associated protein (TCAP)
 CC expression, such as colonic cancer. For example, (I) and (II) may be used
 CC to treat disorders associated with decreased expression by rectifying
 CC mutations or deletions in a patient's genome that affect the activity of
 CC TCAPs by expressing inactive proteins or to supplement the patient's own
 CC production of them. Additionally, (II) may be used to produce the TCAP
 CC proteins, by inserting the nucleic acids into a host cell culturing the
 CC cell to express the protein. (II) and its complementary sequences may
 CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)
 CC and hybridisation assays to detect and quantitate the presence of similar
 CC nucleic acids in samples, and therefore which patients may be in need of
 CC restorative therapy. (I) may also be used as antigens in the production
 CC of antibodies against TCAPs and in assays to identify modulators of TCAP
 CC expression and activity. Anti-(I) antibodies and antagonists may also be
 CC used to down regulate TCAP expression and activity. The anti-(I)
 CC antibodies may also be used as diagnostic agents for detecting the
 CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). AA128460 to AA129512 and AAM24494 to AAM24523 represent
 CC nucleotide and amino acid sequences given in the exemplification of the
 CC present invention
 XX
 SQ Sequence 158 AA;
 XX
 Query Match 100.0%; Score 878; DB 4; Length 158;
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASRSRLLLLSCLAKTGVLDIMRPSCAPGWFYHKSNCYGFRLKRNMSDALEEQS 60
 DB 1 MASRSRLLLLSCLAKTGVLDIMRPSCAPGWFYHKSNCYGFRLKRNMSDALEEQS 60
 QY 61 YGNGAHLASIIISLKEASTIAEYISGYORSQPIWIGLHPQKROQWIDGAMLYRWSWG 120
 DB 61 YGNGAHLASIIISLKEASTIAEYISGYORSQPIWIGLHPQKROQWIDGAMLYRWSWG 120
 QY 121 KSMGNGKHCAMSSNNNFLTWSNCKRQHFLLCKYRP 158
 DB 121 KSMGNGKHCAMSSNNNFLTWSNCKRQHFLLCKYRP 158
 RESULT 8
 AAM24520
 ID AAM24520 standard; protein; 158 AA.
 XX
 AC AAM24520;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE C880P similar amino acid sequence (GENESEQ W37929).
 XX
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
 KW gene therapy; vaccine; colonic cancer.

XX OS Homo sapiens.
XX PN MO200149716-A2.
XX PD 12-JUL-2001.
XX PF 29-DEC-2000; 2000MO-US0355596.
XX PR 30-DEC-1999; 99US-00476296.
XX PR 10-JAN-2000; 2000US-00480321.
XX PR 15-FEB-2000; 2000US-00504629.
XX PR 06-MAR-2000; 2000US-00519444.
XX PR 19-MAY-2000; 2000US-00575251.
XX PR 29-JUN-2000; 2000US-00609448.
XX PR 28-AUG-2000; 2000US-00649811.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stork JA;
XX PI King GE, Wang T, Jiang Y;
XX PD WPI; 2001-441847/47.
XX PT Colon tumor associated proteins and nucleic acids useful for the
XX PT prevention, diagnosis and treatment of colonic cancer.
XX PS Claim 2; Page 469; 472pp; English.
XX CC The present invention describes colon tumour associated proteins (I) and
XX CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
XX CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
XX CC (II) may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate colon tumour associated protein (TCAP)
XX CC expression, such as colonic cancer. For example, (I) and (II) may be used
XX CC to treat disorders associated with decreased expression by rectifying
XX CC mutations or deletions in a patient's genome that affect the activity of
XX CC TCAPs by expressing inactive proteins or to supplement the patients own
XX CC production of them. Additionally, (II) may be used to produce the TCAP
XX CC proteins, by inserting the nucleic acids into a host cell culturing the
XX CC cell to express the protein. (II) and its complementary sequences may
XX CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)
XX CC and hybridisation assays to detect and quantitate the presence of similar
XX CC nucleic acids in samples, and therefore which patients may be in need of
XX CC restorative therapy. (I) may also be used as antigens in the production
XX CC of antibodies against TCAPs and in assays to identify modulators of TCAP
XX CC expression and activity. Anti-(I) antibodies and antagonists may also be
XX CC used to down regulate TCAP expression and activity. The anti-(I)
XX CC antibodies may also be used as diagnostic agents for detecting the
XX CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay
XX CC (ELISA)). AA128460 to AA129512 and AA24494 to AA24523 represent
XX CC nucleotide and amino acid sequences given in the exemplification of the
XX CC present invention
XX SQ Sequence 158 AA;
QY Query Match 100.0%; Score 878; DB 4; Length 158;
Db Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRSMRLLLSLCAKTGVGDIIMRPSCAPGWFYHNSNCGYFRKLRNMSDALEBQS 60
Db 1 MASRSMRLLLSLCAKTGVGDIIMRPSCAPGWFYHNSNCGYFRKLRNMSDALEBQS 60
QY 61 YGNGAHLASITSLKASTIAEYISGYORSOPIMIGLHPQKRQOMQWIDGAMYLTRSMG 120
Db 61 YGNGAHLASITSLKASTIAEYISGYORSOPIMIGLHPQKRQOMQWIDGAMYLTRSMG 120
QY 121 KSMGNGKCAEMSSNNNFLTWSNCECNKRQHFCKYRP 158
Db 121 KSMGNGKCAEMSSNNNFLTWSNCECNKRQHFCKYRP 158

RESULT 9
ID AAM24521
AC AAM24521; standard; protein, 158 AA.
AC AAM24521;
DT 12-OCT-2001 (first entry)
XX DE C880P similar amino acid sequence (GENESEQ W84274).
XX KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
XX KW gene therapy; vaccine; colonic cancer.
XX OS Homo sapiens.
XX PN MO200149716-A2.
XX PD 12-JUL-2001.
XX PF 29-DEC-2000; 2000MO-US0355596.
XX PR 30-DEC-1999; 99US-00476296.
XX PR 10-JAN-2000; 2000US-00480321.
XX PR 15-FEB-2000; 2000US-00504629.
XX PR 06-MAR-2000; 2000US-00519444.
XX PR 19-MAY-2000; 2000US-00575251.
XX PR 29-JUN-2000; 2000US-00609448.
XX PR 28-AUG-2000; 2000US-00649811.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stork JA;
XX PI King GE, Wang T, Jiang Y;
XX PD WPI; 2001-441847/47.
XX PT Colon tumor associated proteins and nucleic acids useful for the
XX PT prevention, diagnosis and treatment of colonic cancer.
XX PS Claim 2; Page 469; 472pp; English.
XX CC The present invention describes colon tumour associated proteins (I) and
XX CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
XX CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
XX CC (II) may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate colon tumour associated protein (TCAP)
XX CC expression, such as colonic cancer. For example, (I) and (II) may be used
XX CC to treat disorders associated with decreased expression by rectifying
XX CC mutations or deletions in a patient's genome that affect the activity of
XX CC TCAPs by expressing inactive proteins or to supplement the patients own
XX CC production of them. Additionally, (II) may be used to produce the TCAP
XX CC proteins, by inserting the nucleic acids into a host cell culturing the
XX CC cell to express the protein. (II) and its complementary sequences may
XX CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)
XX CC and hybridisation assays to detect and quantitate the presence of similar
XX CC nucleic acids in samples, and therefore which patients may be in need of
XX CC restorative therapy. (I) may also be used as antigens in the production
XX CC of antibodies against TCAPs and in assays to identify modulators of TCAP
XX CC expression and activity. Anti-(I) antibodies and antagonists may also be
XX CC used to down regulate TCAP expression and activity. The anti-(I)
XX CC antibodies may also be used as diagnostic agents for detecting the
XX CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay
XX CC (ELISA)). AA128460 to AA129512 and AA24494 to AA24523 represent
XX CC nucleotide and amino acid sequences given in the exemplification of the
XX CC present invention
XX SQ Sequence 158 AA;
QY Query Match 100.0%; Score 878; DB 4; Length 158;
Db Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRSMRLLLSLCAKTGVGDIIMRPSCAPGWFYHNSNCGYFRKLRNMSDALEBQS 60

Db 1 MASSRMLLLLSCLATGVLDIIMRSCAPGWFYHKSNCYGYFRKLRNMSDALEECQS 60
 QY 61 YGNGAHLASTLSLKEASTIAEYISGYORSOPITWGLHDPOKROQOWIDGAMLYYRSMSG 120
 Db 61 YGNGAHLASTLSLKEASTIAEYISGYORSOPITWGLHDPOKROQOWIDGAMLYYRSMSG 120
 QY 121 KSMGNKHCAMSSNNNFLTWSNCKRQHFLLCKYRP 158
 Db 121 KSMGNKHCAMSSNNNFLTWSNCKRQHFLLCKYRP 158

RESULT 10
 AAM24517
 ID AAM24517 standard; protein; 158 AA.
 XX
 AC AAM24517;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE CS1-152 clone predicted amino acid sequence.
 XX
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
 KM gene therapy; vaccine; colonic cancer.
 XX
 OS Homo sapiens.
 XX
 MO200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PE 29-DEC-2000; 2000WO-US035596.
 XX
 PF 30-DEC-1999; 99US-00476296.
 XX
 PR 10-JAN-2000; 2000US-00480321.
 XX
 PR 15-FEB-2000; 2000US-00504629.
 XX
 PR 06-MAR-2000; 2000US-00519444.
 XX
 PR 19-MAY-2000; 2000US-00575251.
 XX
 PR 29-JUN-2000; 2000US-00609448.
 XX
 PR 28-AUG-2000; 2000US-00649811.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 P1 King GE, Wang T, Jiang Y;
 XX
 DR WPI; 2001-441847/47.
 XX
 PT Colon tumor associated proteins and nucleic acids useful for the
 PT prevention, diagnosis and treatment of colonic cancer.
 XX
 PS Claim 2; Page 463; 472pp; English.

The present invention describes colon tumour associated proteins (I) and the polynucleotides (II) that encode them. (I) have cytostatic activity. (I) and (II) can be used in gene therapy and vaccine production. (I) and (II) may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate colon tumour associated protein (TCAP) expression, such as colonic cancer. For example, (I) and (II) may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of TCAPs by expressing inactive proteins or to supplement the patient's own production of them. Additionally, (II) may be used to produce the TCAP proteins, by inserting the nucleic acids into a host cell culturing the cell to express the protein. (II) and its complementary sequences may also be used as DNA probes in diagnostic polymerase chain reaction (PCR) and hybridisation assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. (I) may also be used as antigens in the production of antibodies against TCAPs and in assays to identify modulators of TCAP expression and activity. Anti-(I) antibodies and antagonists may also be used to down regulate TCAP expression and activity. The anti-(I) antibodies may also be used as diagnostic agents for detecting the

CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). AA128460 to AA129512 and AAM24494 to AAM24523 represent
 CC nucleotide and amino acid sequences given in the exemplification of the
 CC present invention
 XX
 SQ Sequence 158 AA;
 Query Match 100.0%; Score 878; DB 4; Length 158;
 Best Local Similarity 100.0%; Pred. No. 6, 8e-83;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASSRMLLLLSCLATGVLDIIMRSCAPGWFYHKSNCYGYFRKLRNMSDALEECQS 60
 Db 1 MASSRMLLLLSCLATGVLDIIMRSCAPGWFYHKSNCYGYFRKLRNMSDALEECQS 60
 QY 61 YGNGAHLASTLSLKEASTIAEYISGYORSOPITWGLHDPOKROQOWIDGAMLYYRSMSG 120
 Db 61 YGNGAHLASTLSLKEASTIAEYISGYORSOPITWGLHDPOKROQOWIDGAMLYYRSMSG 120
 QY 121 KSMGNKHCAMSSNNNFLTWSNCKRQHFLLCKYRP 158
 Db 121 KSMGNKHCAMSSNNNFLTWSNCKRQHFLLCKYRP 158

RESULT 11
 AAM24518
 ID AAM24518 standard; protein; 158 AA.
 XX
 AC AAM24518;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE C880P similar amino acid sequence (GENESBQ W12691).
 XX
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
 KM gene therapy; vaccine; colonic cancer.
 XX
 OS Homo sapiens.
 XX
 MO200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PE 29-DEC-2000; 2000WO-US035596.
 XX
 PF 30-DEC-1999; 99US-00476296.
 XX
 PR 10-JAN-2000; 2000US-00480321.
 XX
 PR 15-FEB-2000; 2000US-00504629.
 XX
 PR 06-MAR-2000; 2000US-00519444.
 XX
 PR 19-MAY-2000; 2000US-00575251.
 XX
 PR 29-JUN-2000; 2000US-00609448.
 XX
 PR 28-AUG-2000; 2000US-00649811.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 P1 King GE, Wang T, Jiang Y;
 XX
 DR WPI; 2001-441847/47.
 XX
 PT Colon tumor associated proteins and nucleic acids useful for the
 PT prevention, diagnosis and treatment of colonic cancer.
 XX
 PS Claim 2; Page 467; 472pp; English.

The present invention describes colon tumour associated proteins (I) and the polynucleotides (II) that encode them. (I) have cytostatic activity. (I) and (II) can be used in gene therapy and vaccine production. (I) and (II) may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate colon tumour associated protein (TCAP) expression, such as colonic cancer. For example, (I) and (II) may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of

CC TCAPs by expressing inactive proteins or to supplement the patients own
CC production of them. Additionally, (ii) may be used to produce the TCAP
CC proteins, by inserting the nucleic acids into a host cell culturing the
CC cell to express the protein. (ii) and its complementary sequences may
CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)
CC and hybridisation assays to detect and quantitate the presence of similar
CC nucleic acids in samples, and therefore which patients may be in need of
CC restorative therapy. (ii) may also be used as antigens in the production
CC of antibodies against TCAPs and in assays to identify modulators of TCAP
CC expression and activity. Anti-(i) antibodies and antagonists may also be
CC used to down regulate TCAP expression and activity. The anti-(ii)
CC antibodies may also be used as diagnostic agents for detecting the
CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay
CC (ELISA)). AA128460 to AA129512 and AA24494 to AA24523 represent
CC nucleotide and amino acid sequences given in the exemplification of the
CC present invention

SQ Sequence 158 AA;

Query Match 100.0%; Score 878; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLILLLISCLAKTGVLDIIMRSPCAPGMFYHKSNCYGYRKLKRNMSDALEECOS 60
DB 1 MASRSMRLILLLISCLAKTGVLDIIMRSPCAPGMFYHKSNCYGYRKLKRNMSDALEECOS 60
QY 61 YNGAHLASITSLKEASTIAEYISGYORSOPITWIGLHDPKROQOWIDGAMYLVSWSG 120
DB 61 YNGAHLASITSLKEASTIAEYISGYORSOPITWIGLHDPKROQOWIDGAMYLVSWSG 120
QY 121 KSMGNKIKCAEWSNNNFLTWSNNECNKRQHFCKYRP 158
DB 121 KSMGNKIKCAEWSNNNFLTWSNNECNKRQHFCKYRP 158

RESULT 12

AAE29829 ID AAE29829 standard; protein; 158 AA.

AC AAE29829;

DT 24-FEB-2003 (first entry)

DE Human REG-like protein (RELp).

KW Human; REG-like protein; RELp; tumour; cancer; therapy; chromosome 1.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT Protein /label=Signal_peptide

FT /note="human mature/REG-like protein"

PN EP1241269-A2.

PD 18-SEP-2002.

PP 15-MAR-2002; 2002EP-00251876.

PR 16-MAR-2001; 2001US-0276414P.

PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.

PI Helsinki M;

DR WPI, 2002-684095/74.

DR N-PSDB; AAD47239.

PT Detecting the presence of a tumor comprises detecting the concentration
PT of a Reg like Protein or the presence or quantity of a nucleic acid

PT encoding it.
XX
PS Claim 1; Page 13-14; 26pp; English.

XX The invention relates to a method for detecting REG-like protein (RELp)
CC and its nucleic acid sequence. The method is useful for detecting the
CC presence of a tumour. Kits comprising an antibody specific for RELp and
CC reagents for detecting the antibody, or a nucleic acid complementary to a
CC portion of a nucleic acid encoding RELp, are useful for identifying the
CC presence of cancer, characterise the cancer, or monitor the course of
CC treatment of cancer. The present sequence is human RELp protein used to
CC illustrate the method of the invention. Human RELp gene is located at
CC chromosome 1

SQ Sequence 158 AA;

Query Match 100.0%; Score 878; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLILLLISCLAKTGVLDIIMRSPCAPGMFYHKSNCYGYRKLKRNMSDALEECOS 60
DB 1 MASRSMRLILLLISCLAKTGVLDIIMRSPCAPGMFYHKSNCYGYRKLKRNMSDALEECOS 60
QY 61 YNGAHLASITSLKEASTIAEYISGYORSOPITWIGLHDPKROQOWIDGAMYLVSWSG 120
DB 61 YNGAHLASITSLKEASTIAEYISGYORSOPITWIGLHDPKROQOWIDGAMYLVSWSG 120
QY 121 KSMGNKIKCAEWSNNNFLTWSNNECNKRQHFCKYRP 158
DB 121 KSMGNKIKCAEWSNNNFLTWSNNECNKRQHFCKYRP 158

RESULT 13

ABB78993 ID ABB78993 standard; protein; 158 AA.

AC ABB78993;

DT 02-AUG-2002 (first entry)

DE Human Reg IV protein sequence SEQ ID NO:4471.

KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;

KW genetic analysis; diagnostic; antisense therapy.

XX Homo sapiens.

XX WO200229086-A2.

XX 11-APR-2002.

PP 02-OCT-2001; 2001WO-US030732.

PR 02-OCT-2000; 2000US-0237271P.

PA (PARB) BAYER CORP.

PI Burgess C, Aetle JH, Carroll E, Catino TV, Dwivedi P, Molino GA,

PI Thiglingam A, Lewis ME,

DR WPI, 2002-426115/45.

DR N-PSDB; ABO60776.

PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.

PS Claim 5; Fig 3; 796pp; English.

CC ABO6306 to ABO60787 represent isolated nucleic acids (i) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABO60776 to ABO60787 nucleic acid sequences. (i) can be

CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (1) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridizes to (1) in a cell. A probe/primer derived
CC from (1) can be used for determining the presence of a nucleic acid which
CC hybridizes to (1), and for determining the phenotype of cells in a sample
CC of cells from a patient. (1) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (1) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
CC
XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRSPCAPGWFYHNSCYGFRKLRNMSDALEECOS 60
DB 1 MASRSMRLLLLSCLAKTGVLDIIMRSPCAPGWFYHNSCYGFRKLRNMSDALEECOS 60
QY 61 YNGNAHLASISLKEASTIAEYISGYORSQPIWIGLHPQKRQOWIDGAMLYRSMWG 120
DB 61 YNGNAHLASISLKEASTIAEYISGYORSQPIWIGLHPQKRQOWIDGAMLYRSMWG 120
QY 121 KSMGNKHCAMSSNNNFLTWSNECNKRQHFCKYRP 158
DB 121 KSMGNKHCAMSSNNNFLTWSNECNKRQHFCKYRP 158

RESULT 14
ABRS551
ID ABRS551 standard; protein; 158 AA.

AC ABR58551;

DT 09-JUN-2003 (first entry)

DE Human cancer related protein SEQ ID NO:208.

KW Human; cancer; diagnosis; screening; modulator; leukemia; ischemia;
KW heart disease; atherosclerosis; endometriosits.

OS Homo sapiens.

PN WO2003025138-A2.

PD 27-MAR-2003.

PF 17-SEP-2002; 2002WO-US029560.

PR 17-SEP-2001; 2001US-0323469P.

PR 20-SEP-2001; 2001US-0323887P.

PR 13-NOV-2001; 2001US-0350666P.

PR 08-FEB-2002; 2002US-0355145P.

PR 08-FEB-2002; 2002US-0355257P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

PI Afer D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;

PI Zlotnick A;

DR WPI: 2003-354600/33.

DR N-PSDB; ACC72672.

XX New genes that are up-regulated or down-regulated in cancers, useful as
XX markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
XX therapeutic targets for screening drugs for treating these diseases.

PS Claim 12; Page 737; 767pp; English.

XX The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC of the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischemia, heart diseases,
CC atherosclerosis and endometriosits. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
CC
XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRSPCAPGWFYHNSCYGFRKLRNMSDALEECOS 60
DB 1 MASRSMRLLLLSCLAKTGVLDIIMRSPCAPGWFYHNSCYGFRKLRNMSDALEECOS 60
QY 61 YNGNAHLASISLKEASTIAEYISGYORSQPIWIGLHPQKRQOWIDGAMLYRSMWG 120
DB 61 YNGNAHLASISLKEASTIAEYISGYORSQPIWIGLHPQKRQOWIDGAMLYRSMWG 120
QY 121 KSMGNKHCAMSSNNNFLTWSNECNKRQHFCKYRP 158
DB 121 KSMGNKHCAMSSNNNFLTWSNECNKRQHFCKYRP 158

RESULT 15

ABPS6022
ID ABPS6022 standard; protein; 158 AA.

AC ABPS6022;

DT 26-FEB-2003 (first entry)

DE Human REG-1-like protein (REL1P) SEQ ID NO:2.

KW Human; REG-1-like protein; REL1P; immunoglobulin derived protein; Ig;
KW immunoglobulin; cytosolic; Ig agonist; immunoglobulin agonist; cancer;
KW protein therapy; REL1P human Ig derived protein; chromosome 1p12-13.1.

OS Homo sapiens.

PN WO200274916-A2.

PD 26-SEP-2002.

PF 14-MAR-2002; 2002WO-US007945.

PR 16-MAR-2001; 2001US-0276305P.

XX (CENZ) CENTOCOR INC.

PI Key Location/Qualifiers

PI Peptide 1..26

PI Protein 27..158

PI /label= REL1P

XX Heiskala M;

DR WPI; 2003-103204/09.

DR N-PSDB; AB221635.

PT New isolated REG-like protein (RELp) human immunoglobulin derived protein or specified portion or variant, useful for preventing or treating a RELp protein mediated condition or malignant condition, e.g. cancer.

PS Claim 1; Fig 2; 101pp; English.

CC The present sequence represents a new isolated REG-like protein (RELp) human immunoglobulin (Ig) derived protein. RELp comprises: (a) a human variable and constant region; or (b) an isolated human Ig derived protein or specified portion or variant encoded by a nucleic acid. RELp has cytostatic activity and can be used as an Ig agonist and in protein therapy. The RELp human Ig derived protein or a specified portion or variant can be used for preventing or treating a RELp protein mediated condition, malignant condition or disease condition, e.g. cancer. The nucleic acids can be used in producing RELp Ig derived protein. The human RELp protein of the present invention is located to chromosome 1p12-13.1

CC Sequence 158 AA;

Query Match 100.0%; Score 878; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.8e-83;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MASRSMRLLLLSCLAKTGVLDITMRPSCAPGFYHKSNCYGYFRKLRNMSDALEECOS 60
DB 1 MASRSMRLLLLSCLAKTGVLDITMRPSCAPGFYHKSNCYGYFRKLRNMSDALEECOS 60
OY 61 YGNGAHLASTLSLKEASTIAEYISGYORSQPIWIGLHDPQKROQWQWIDGAMTYLYRSWSG 120
DB 61 YGNGAHLASTLSLKEASTIAEYISGYORSQPIWIGLHDPQKROQWQWIDGAMTYLYRSWSG 120
OY 121 KSMGNGKCAEMSSNNFLTWSSNECNKRQHFLLCKYRP 158
DB 121 KSMGNGKCAEMSSNNFLTWSSNECNKRQHFLLCKYRP 158

Search completed: February 11, 2005, 23:29:37
Job time : 169 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 23:16:55 ; Search time 177 Seconds
(without alignments)
457.110 Million cell updates/sec

Title: US-09-525-041-2

Sequence: 1 MASRSMRLLLLSCLAKGV.....LTWSSNECKRQHFLCKTRP 158
Perfect score: 878

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	878	100.0	158	2	09BYZ8
2	600.5	68.4	157	2	09BDG5
3	594.5	67.7	157	2	09DB58
4	593.5	67.6	157	2	068AX7
5	544	62.0	113	2	08NER7
6	438.5	49.9	160	2	07SE75
7	307.5	35.0	134	2	08NER6
8	260.5	29.7	132	1	STR1_STRCA
9	258.5	29.4	132	1	ACAL_ANSAN
10	254.5	29.0	165	1	LIT1_MOUSE
11	254.5	29.0	165	1	LITR_RAT
12	247	28.1	142	1	STR2_STRCA
13	246.5	28.1	174	1	PAP3_MOUSE
14	245	27.9	164	2	06TR56
15	243.5	27.7	158	2	06QK33
16	242	27.6	175	1	PAP2_MOUSE
17	241.5	27.5	142	1	OC17_CHICK
18	240	27.3	166	1	LITR_HUMAN
19	237.5	27.0	158	2	07ARQ1
20	237	27.0	135	1	LECG_BOTUR
21	236.5	26.9	173	1	LIT2_MOUSE
22	236	26.9	176	2	07T228
23	235	26.8	126	2	08CE69
24	234	26.7	135	1	LECG_BITAR
25	233	26.5	135	1	LECG_CROAT
26	231	26.3	174	1	PAP3_RAT
27	231	26.3	175	1	LITR_BOVIN
28	230.5	26.3	154	2	08UIV9
29	229	26.1	135	1	LECG_LACST
30	228.5	26.0	174	1	PAP2_RAT
31	227.5	25.9	146	2	09CVF4

32	227	25.9	158	2	09OWI7	09OWI7 bungarus fa
33	227	25.9	158	2	09OWI8	09OWI8 bungarus fa
34	225.5	25.7	146	2	06X5S1	06X5S1 echis pyram
35	225.5	25.7	146	2	06X5S4	06X5S4 echis carin
36	225.5	25.7	166	1	LITR_HUMAN	P05451 homo sapien
37	222.5	25.3	172	1	LECG_PLEMA	002958 pleurodeles
38	222	25.3	157	2	06T7B5	06T7B5 bittis gabon
39	221.5	25.2	155	2	08TIV8	08TIV8 bittis gabon
40	221	25.2	175	1	PAP1_HUMAN	006141 homo sapien
41	219.5	25.0	148	2	06X5S3	06X5S3 echis pyram
42	219.5	25.0	148	2	06X5S7	06X5S7 echis ocell
43	219.5	25.0	148	2	06X5S9	06X5S9 echis carin
44	219.5	25.0	148	2	06X5T1	06X5T1 bittis ariet
45	218.5	24.9	154	2	07T2Q0	07T2Q0 echis multt

ALIGNMENTS

RESULT 1
09BYZ8 PRELIMINARY; PRT; 158 AA.
AC 09BYZ8;
DT 01-JUN-2001 (JEMBLrel. 17, Created)
DT 01-JUN-2001 (JEMBLrel. 17, Last sequence update)
DT 25-OCT-2004 (JEMBLrel. 28, Last annotation update)
DE Regenerating gene type IV precursor (REG-like protein) (Regenerating
DE islet-derived family, member 4) (Gastrointestinal secretory protein
DE GISP).
GN Name=REG4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21210973; PubMed=11311942; DOI=10.1016/S0167-4781(00)00284-0;
RX Hartupree J.C., Zhang H., Bonaldo M.F., Soares M.B., Dieckgraber B.K.;
RT "Isolation and characterization of a cDNA encoding a novel member of
RL the human regenerating protein family: Reg IV(1)."
RL Biochim. Biophys. Acta 1518:287-293(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Karamanen M., Heiskala K., Heiskala M., Andersson L.C.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Karamanen M., Heiskala K., Heiskala M., Andersson L.C.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RX Klausner R.D., Fellingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Dischenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Caavaant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boesak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Huiyik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Straubeberg R.;
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Lin W.-C.;
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY007243; AAG02562.1; -
 DR EMBL; AF345934; AAK5986.1; -
 DR EMBL; AY126670; AAM95598.1; -
 DR EMBL; BC017089; AAK17089.1; -
 DR EMBL; AF254415; AAK48435.1; -
 DR HSSP; P22030; IITK.
 DR Genew; HGNC:22977; REG4.
 DR GO; GO:0005529; F: sugar binding; IEA.
 DR InterPro; IPR01304; Lectin_C.
 DR InterPro; IPR03990; Pancreatis_ac.
 DR Pfam; PF00059; Lectin_C_1.
 DR PRINTS; PR01504; PNCEARITSA.
 DR SMART; SM00034; CLECT_1.
 DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
 DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 158 AA; 18230 MW; 7308849CBBD6E93E CRC64;
 Query Match 100.0%; Score 878; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 2,5e-78;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLTGDIIMRPSGAPGWFYHKSNCYGRKRLNMSDALEECQS 60
 DB 1 MASRSMRLLLLSCLAKTGVLTGDIIMRPSGAPGWFYHKSNCYGRKRLNMSDALEECQS 60
 QY YGNNGHLLSLTSLKASRTAETISGYORSOPITGLHPQKQOWIDGMYLYTRMSG 120
 DB YGNNGHLLSLTSLKASRTAETISGYORSOPITGLHPQKQOWIDGMYLYTRMSG 120
 QY 121 KSMGNGKCAEWSNNFLLTSSNCCNCRKQHFLLCKYRP 158
 DB 121 KSMGNGKCAEWSNNFLLTSSNCCNCRKQHFLLCKYRP 158

RESULT 2
 Q9D8G5 PRELIMINARY; PRT; 157 AA.
 ID Q9D8G5
 AC Q9D8G5
 DT 01-UN-2001 (TREMBlrel. 17, Created)
 DT 01-UN-2001 (TREMBlrel. 17, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Mus musculus adult male small intestine cDNA, RIKEN full-length
 DE enriched library, clone:2010002L15 product:REGENERATING GENE TYPE IV,
 DE full insert sequence (Reg4 protein).
 GN Name=Reg4;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Small intestine;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carrincci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Mech. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Small intestine;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Small intestine;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Small intestine;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carrincci P., Shihata Y., Hayatsu N., Sugahara Y., Shihata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subcloning of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Small intestine;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shihata K., Itoh M., Alzawa K., Nagaoaka S., Sasaki N., Carrincci P.,
 RA Konno H., Akiyama J., Nishi K., Kitahara T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaishige K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanuki M.,
 RA Yoneda Y., Ishikawa T., Osawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Small intestine;
 RA Adachi J., Alzawa K., Akihara S., Akimura T., Arai A., Aono H.,
 RA Arikawa T., Bono H., Carrincci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hasegaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imetani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato C.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numasaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shihata K., Shihata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schutler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hesteh L.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stappleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrincci P., Prange C.,
 RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wortley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smallos D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Warr M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

RN [8]
RC SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Colon;
RA Struhsberg R.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK008049; BAB25429.1; -
DR EMBL; BC019465; AAH19465.1; -
DR HSSP; 006141; IUV0.
DR MGD; MGI:1914959; Reg4.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancratic_ac.
DR Pfam; PF00059; Lectin_C_1.
DR PRINTS; PR01504; PNCEATITSP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 157 AA; 18398 MW; F3981722BBD83968 CRC64;

Query Match 68.4%; Score 600.5; DB 2; Length 157;
Best Local Similarity 66.2%; Pred. No. 4,8e-51;
Matches 104; Conservative 25; Mismatches 27; Indels 1; Gaps 1;

QY 1 MASRSRMLLLLSCTAKTGVLDIMRPSCAPGWFYHNSCYGRKJLNMSDALEECOS 60
DB 1 MASRGRVRLLLLSWVAGPEVLSL-ILRPSCAPGWFYRSHCYGRKJLNMSHALEECOS 59
QY 61 YNGNHLASIIISLKEASTIAEYISGYORSOPWIGLHDPQRQOWIDGAMYLYRSWSG 120
DB 60 YNGSHLASVNLNQKASVYSKYITGQRLPWIGLHDPQRKQLQWMTDGSNTLYRRWNP 119
QY 121 KSMGNKCAEMSSNNFLTWSNECNKRQHFLLCYR 157
DB 120 RTKSEARHCAMNPDKFLTNKNGCANRQHFLLCYR 156

RESULT 3
Q9D858 PRELIMINARY; PRT; 157 AA.
ID Q9D858
AC 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DR 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Mus musculus adult male small intestine cDNA, Riken full-length
DE enriched library, clone:2010204K21 product:REGENERATING GENE TYPE IV,
DE full insert sequence.
GN Name=Reg4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Small intestine;
RA MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999);
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Small intestine;
RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Small intestine;
RA The RIKEN Consortium;
RA "Analysis of the mouse transcriptome based on functional annotation of
RA 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Small intestine;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Small intestine;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Atzawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama U., Nishi K., Kitasuna T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Small intestine;
RA Adachi J., Atzawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukuishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imocani K., Ishii Y., Itoh M., Izawa M., Kaubawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Maruyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Saki C., Saki K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK008438; BAB25669.1; -
DR HSSP; 006141; IUV0.
DR MGD; MGI:1914959; Reg4.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancratic_ac.
DR Pfam; PF00059; Lectin_C_1.
DR PRINTS; PR01504; PNCEATITSP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 157 AA; 18474 MW; FD96F36CFB989368 CRC64;

Query Match 67.7%; Score 594.5; DB 2; Length 157;
Best Local Similarity 65.6%; Pred. No. 1.9e-50;
Matches 103; Conservative 25; Mismatches 28; Indels 1; Gaps 1;

QY 1 MASRSRMLLLLSCTAKTGVLDIMRPSCAPGWFYHNSCYGRKJLNMSDALEECOS 60
DB 1 MAYKGVRLLLLSWVAGPEVLSL-ILRPSCAPGWFYRSHCYGRKJLNMSHALEECOS 59
QY 61 YNGNHLASIIISLKEASTIAEYISGYORSOPWIGLHDPQRQOWIDGAMYLYRSWSG 120
DB 60 YNGSHLASVNLNQKASVYSKYITGQRLPWIGLHDPQRKQLQWMTDGSNTLYRRWNP 119
QY 121 KSMGNKCAEMSSNNFLTWSNECNKRQHFLLCYR 157
DB 120 RTKSEARHCAMNPDKFLTNKNGCANRQHFLLCYR 156

RESULT 4
Q68AX7 PRELIMINARY; PRT; 157 AA.
ID Q68AX7
AC Q68AX7;
DT 25-OCT-2004 (TREMblrel. 28, Created)
DT 25-OCT-2004 (TREMblrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)

```

DE Regenerating islet-derived family member 4.
GN Name=Reg4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistar; TISSUE=ileum;
RA Naitaka K., Murakami K., Fukushima M., Kiyama H.;
RT "Differential regulation of Reg family member expression after
RT peripheral nerve injury";
RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB164049; BAD38673.1; -;
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatls_ac.
DR Pfam; PF00059; Lectin_C.1.
DR PRINTS; PRO1504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 157 AA; 18265 MW; 558129FB10BA4D1D CRC64;

Query Match 67.6%; Score 593.5; DB 2; Length 157;
Best Local Similarity 66.5%; Pred. No. 2.3e-50;
Matches 105; Conservative 21; Mismatches 31; Indels 1; Gaps 1;
Qy 1 MASRMRLLLSLCTAKTGVIGDIIIRPSCAPGWFYHKSNCYGFRKLRNMSDALEECOS 60
Db 1 MASKKVRLLLSLWAGPEVLSD-ILRPSCASGMVNRSHCGYFRKLRNMSHALEECOS 59
Qy 61 YGNGAHLSILSLKEASTIAEYISGYORSOPITWIGLHPORQOQWOWIDGAMVLYRSMSG 120
Db 60 YGNGSHLASLVNPKKASVISKITYGQSLPWTIGLHPQKNASQWIDGSTNQVRPMPSP 119
Qy 121 KSMGNGKCAEENSSNNNFLTSSNECNKRQHFLECYRP 158
Db 120 RTKSEARHCTEMNPDKFLTWNNKCTCRQHFLECYRP 157

RESULT 5
Q8NER7 PRELIMINARY; PRT; 113 AA.
ID Q8NER7;
AC Q8NER7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE REG-like protein splice variant 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Kamatani M., Heiskala K., Heiskala M., Andersson L.C.;
RA Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY126671; AAM95599.1; -;
DR HSSP; P21963; IJZN.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatls_ac.
DR Pfam; PF00059; Lectin_C.1.
DR PRINTS; PRO1504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 113 AA; 12832 MW; A2E9DPLA729C78DA CRC64;

Query Match 62.0%; Score 544; DB 2; Length 113;
Best Local Similarity 99.0%; Pred. No. 1.2e-45;
Matches 101; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MASRMRLLLSLCTAKTGVIGDIIIRPSCAPGWFYHKSNCYGFRKLRNMSDALEECOS 60
Db 1 MASRMRLLLSLCTAKTGVIGDIIIRPSCAPGWFYHKSNCYGFRKLRNMSDALEECOS 60

Qy 61 YGNGAHLSILSLKEASTIAEYISGYORSOPITWIGLHPORQ 102
Db 61 YGNGAHLSILSLKEASTIAEYISGYORSOPITWIGLHPORQ 102

RESULT 6
Q78275 PRELIMINARY; PRT; 160 AA.
ID Q78275;
AC Q78275;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE MG664513 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OX Xenopodinae; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=whole;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins B., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ueda T.B., Toshimiyuki S., Carantini P., Prange C.,
RA Bosa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimmond J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=whole;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC053817; AAH53817.1; -;
DR HSSP; P22897; IREG.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_C.1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 160 AA; 18330 MW; 6A5502F24689179A CRC64;

Query Match 49.9%; Score 438.5; DB 2; Length 160;
Best Local Similarity 48.7%; Pred. No. 4.1e-35;
Matches 75; Conservative 35; Mismatches 41; Indels 3; Gaps 2;
Qy 8 LLLLSLCTAKTGVIGDIIIRPSCAPGWFYHKSNCYGFRKLRNMSDALEECOSYGNGAHL 67
Db 8 LLLLSLCTAKTGVIGDIIIRPSCAPGWFYHKSNCYGFRKLRNMSDALEECOSYGNGAHL 67

```

Db      7 LLLPGALAVSNVLEAAYRSSCPNMGPFYKANCYGFYRPLSMABEAYDCQAYGHAML 66
QY      68 ASTLSKEASTAEVYSGVRSOPIMVIGLHPKROQWQMDGAMLYNSW-SGKMGCN 126
Db      67 ASTLSDAEADVASHSHSAVQKNPVMIGLHDPFQNRKMKNDGSMYNSWLAGDPDNYN 126
QY      127 --KHCAEMSSNNFLTWSSNECNKROHFLCKYAP 158
Db      127 SAEYGEISCKEGFVAKMNSCKEYKQYCKYCP 160

RESULT 7
Q8MER6  PRELIMINARY; PRT; 134 AA.
ID      08NER6;
AC      08NER6;
DT      01-OCT-2002 (TEMBLrel. 22, Created)
DT      01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DE      REG-like protein splice variant 2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Kamarainen M., Heiskala K., Heiskala M., Andersson L.C.;
RL      Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; A126672; AM93960-1;
SQ      SEQUENCE 134 AA; 14993 MW; E0E5AD9B6A53EB5 CRC64;

Query Match      35.0%; Score 307.5; DB 2; Length 134;
Best local Similarity 55.9%; Pred. No. 2.6e-22;
Matches 62; Conservative 9; Mismatches 9; Indels 31; Gaps 3;

QY      1 MASRSRRLLLLSCLAKTGYLGDIIKRPSCAPGFYKNSCYGFKRLNWSDAEL---- 56
Db      1 MASRSRRLLLLSCLAKTGYLGDIIKRPSCAPGFYKNSCYGFKRLNWSDAELVNLL 60

QY      57 -----ECOSYNGAHLASILSKEASTAEVYSGVRSOPIMV 94
Db      61 PAMPGLSRAKQDPEPQ-----ISPDGSSV--LPGHYEKPIML 97

RESULT 8
STR1_STRCA
ID      STR1_STRCA STANDARD; PRT; 132 AA.
AC      P83514;
DT      05-JUL-2004 (Rel. 44, Created)
DT      05-JUL-2004 (Rel. 44, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Struthio camelus (Ostrich).
OS      Struthio camelus (Ostrich).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
OC      Struthio.
NCBI_TaxID=8801;
RN      [1]
RP      SEQUENCE, SUBCELLULAR LOCATION, AND MASS SPECTROMETRY.
RC      TISSUE=Eggshell matrix;
RX      PubMed=14726203; DOI=10.1016/j.bbapap.2003.09.006;
RA      Mann K., Siedler F.;
RT      "Ostrich (Struthio camelus) eggshell matrix contains two different C-
RT      type lectin-like proteins. Isolation, amino acid sequence, and
RT      posttranslational modifications."
RL      Biochim. Biophys. Acta 1696:41-50(2004).
CC      -1- SUBCELLULAR LOCATION: Eggshell matrix.
CC      -1- MASS SPECTROMETRY: MW=15343.2; MW_ERR=4; METHOD=Electrospray;
CC      RANGE=1-133; NOTE=Ref.1.
CC      -1- SIMILARITY: Contains 1 C-type lectin family domain.
DR      InterPro; IPR002353; AntifreezeZell.
DR      InterPro; IPR001304; Lectin_C.
DR      InterPro; IPR003990; Pancreatit_ac.
PFam; PF00059; Lectin_c; 1.

```

```

DR      PRINTS; PR00356; ANTIFREEZEII.
DR      PRINTS; PR01504; PNCREATITSAF.
DR      SMART; SM00034; CLECT. 1.
DR      PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR      PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW      Direct protein sequencing; Lectin.
FT      DOMAIN 10 129 C-type lectin.
FT      DISULFID 3 14 By similarity.
FT      DISULFID 31 126 By similarity.
FT      DISULFID 103 120 By similarity.
SQ      SEQUENCE 132 AA; 15353 MW; F7BDIDF2990B2945 CRC64;

Query Match      29.7%; Score 260.5; DB 1; Length 132;
Best local Similarity 38.8%; Pred. No. 1.1e-17;
Matches 50; Conservative 20; Mismatches 54; Indels 5; Gaps 2;

QY      30 CAPGFYKNSCYGFKRLNWSDAELFCQSYNGAHLASILSKEASTAEVYSGY--- 86
Db      3 CPKGMIDFRGNCYGFYRELPMKRAEAWCRSIRAGHILASIHTEBHRAIKFIISQYHHG 62

QY      87 QRSQPIWIGLHPQKROQWQMDGAMLYRSMGSKMGKCAEMSSNNFLTWSSNRC 146
Db      63 EEEEDVWIGLF--RMSVWAWIDGSKKHYSAALDDDPKXKCAVLDESSGFLSMWNSDC 120

QY      147 NKROHFLCK 155
Db      121 GERNAFICK 129

RESULT 9
ACAL_ANSAN
ID      ACAL_ANSAN STANDARD; PRT; 132 AA.
AC      P83300;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Ansoalcain.
OS      Anser anser anser (Western graylag goose).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
OX      NCBI_TaxID=8844;
RN      [1]
RP      SEQUENCE, FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND MASS
RP      SPECTROMETRY.
RC      TISSUE=Eggshell matrix;
RX      MEDLINE=22439773; PubMed=12431998; DOI=10.1074/jbc.M201518200;
RA      Lakshminarayanan R., Valliyaveetil S., Rao V.S., Kini R.M.;
RT      "Purification, characterization, and in vitro mineralization studies
RT      of a novel goose eggshell matrix protein, ansocalcin."
RL      J. Biol. Chem. 278:2928-2936(2003).
CC      -1- FUNCTION: Induces spherical aggregates of calcite crystals in
CC      vitro. Believed to play an active role in the eggshell
CC      calcification.
CC      -1- SUBUNIT: Homodimer or homotrimer.
CC      -1- SUBCELLULAR LOCATION: Eggshell matrix. May be present in the shell
CC      glands on the walls of oviduct and incorporated into the shell
CC      structure during its formation.
CC      -1- MASS SPECTROMETRY: MW=15342; METHOD=Electrospray; RANGE=1-132;
CC      NOTE=Ref.1.
CC      -1- SIMILARITY: Contains 1 C-type lectin family domain.
DR      HSP; P23806; L134.
DR      InterPro; IPR002353; AntifreezeZell.
DR      InterPro; IPR001304; Lectin_C.
DR      InterPro; IPR003990; Pancreatit_ac.
PFam; PF00059; Lectin_C; 1.
DR      PRINTS; PR00356; ANTIFREEZEII.
DR      PRINTS; PR01504; PNCREATITSAF.
DR      SMART; SM00034; CLECT. 1.
DR      PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR      PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW      Direct protein sequencing; Lectin.
FT      DOMAIN 1 132 C-type lectin.
FT      DISULFID 3 14 By similarity.

```

```

FT DISULFID 31 128 By similarity.
FT DISULFID 103 120 By similarity.
SQ SEQUENCE 132 AA; 15347 MW; 36CE42BA57226B7 CRC64;

Query Match 29.4%; Score 258.5; DB 1; Length 132;
Best Local Similarity 34.3%; Pred. No. 1.7e-17;
Matches 46; Conservative 25; Mismatches 50; Indels 13; Gaps 3;

QY 30 CAPGFHYKSNQCYFRKLRNMSDAELPGQSYGNGAHILASITLSKEASTIAEYISGYORS 89
DB 3 CPKGFHLDGRSGCYGFGGELTRKAEAMCKVIHACHLASHSPSEHAAVAFKFPQR 62
QY 90 QP--IWIGLHDPKROQMO---WIDGAMYLRSWSGKMGNGHCAEMSSNNFLTWS 142
DB 63 EEDNDVWVIGLH-----HMNQARVWIDSGKKRYSAMDDDELPRGKYCTVIBSSSGFME 116
QY 143 SNECKNRQHFUCKY 156
DB 117 DNACSEBNPFCVKY 130

RESULT 10
LITL MOUSE
LITL MOUSE STANDARD; PRT; 165 AA.
AC P43137;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lithostathine 1 precursor (pancreatic stone protein 1) (PSP)
DE Pancreatic thread protein 1 (PTP) (Islet of Langerhans regenerating
DE protein 1) (REG 1).
GN Name=Reg1;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=93340209; PubMed=8340418;
RA Ueno M., Yonekura H., Nakagawara K.-I., Watanabe T., Miyashita H.,
RA Morizumi S., Okamoto H., Itoh T., Teraoka H.;
RA "Structure, chromosomal localization, and expression of mouse reg
RA genes, reg I and reg II. A novel type of reg gene, reg II, exists in
RA the mouse genome.";
RL J. Biol. Chem. 268:15974-15982(1993).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Struhsberg R.L., Felsing E.A., Grouse I.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Alteschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Cantinchi P., Prange C.,
RA Rahn S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.V., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Sherchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.J., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Might act as an inhibitor of spontaneous calcium
CC carbonate precipitation.
CC -1- TISSUE SPECIFICITY: Expressed only in regenerating islets and

```

```

CC normal exocrine pancreas, but not in normal pancreatic islets.
CC Expressed strongly in pancreas, moderately in gallbladder, and
CC weakly in liver.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; D14010; BAA0311.1; -.
DR EMBL; BC028761; AA028761.1; -.
DR PIR; A47148; A47148.
DR HSSP; P05451; LITL.
DR MGD; MGI:97895; Reg1.
DR InterPro; IPR002353; Antifreeze1.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatic_ac.
DR Pfam; PF00059; Lectin_C_1.
DR PRINTS; PR00356; ANTIFREEZE1.
DR PRINTS; PR01504; PNCRAVITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KM Glycoprotein; Lectin; Signal.
FT SIGNAL 1 21 By similarity.
FT CHAIN 22 165 Lithostathine 1.
FT DOMAIN 33 163 C-type lectin.
FT DISULFID 35 46 By similarity.
FT DISULFID 63 161 By similarity.
FT DISULFID 136 153 By similarity.
FT CARBOHYD 129 129 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 165 AA; 18518 MW; 2950174AF5D66BA CRC64;

Query Match 29.0%; Score 254.5; DB 1; Length 165;
Best Local Similarity 31.1%; Pred. No. 5.2e-17;
Matches 52; Conservative 35; Mismatches 51; Indels 29; Gaps 7;

QY 10 LLSCLAATGYLADIIKRP-----SCAPGFHYKSNQCYFRKLRNMSDA 54
DB 8 ILLSCL-----IVLSPSQGEAEEDLPPEARISCEGSAVASYCYFTEDRLTWADA 59
QY 55 ELEQSYGNGAHILASITLSKEASTIAEYI--SGYRQRPVITGLHDPKROQMOIWIGAM 112
DB 60 DLPCQNMNSG-YLVSVLSQAEGNFVASLIKESGTTDAN-VWTGLHDPKRRMRHWSGSL 117
QY 113 YLYRSGW-SGKSGMGNK-HCAEMSSNNFLTWSNECKNRQHFUCKYR 157
DB 118 FLYRSMATGPNSSNRGTCVLSYNTGYKKMKDNCDAQYFVCKFK 164

RESULT 11
LITL RAT
LITL RAT STANDARD; PRT; 165 AA.
AC P10758;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lithostathine precursor (pancreatic stone protein) (PSP) (pancreatic
DE thread protein) (PTP) (Islet of Langerhans regenerating protein) (REG)
DE (Islet cells regeneration factor) (ICRF).
GN Name=Reg1; Synonyms=Reg;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=91093273; PubMed=1985964;
RA Rouquier S., Verdier J.M., Iovanna J., Dagorn J.-C., Giorgi D.;

```


RT	"Rat pancreatic stone protein messenger RNA. Abundant expression in mature exocrine cells, regulation by food content, and sequence identity with the endocrine reg transcript.";
RT	J. Biol. Chem. 266:786-791(1991).
RN	(2)
RP	SEQUENCE FROM N.A.
RX	MEDLINE=86115343; PubMed=2963000;
RA	Terazono K., Yamamoto H., Takasawa S., Shiga K., Yonemura Y., Tochino Y., Okamoto H.;
RA	"A novel gene activated in regenerating islets.";
RL	J. Biol. Chem. 263:2111-2114(1988).
RN	[3]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=93326645; PubMed=7916640; DOI=10.1016/0167-4781(93)90100-R;
RA	Ducret N.J., Frigerio J.M., Dagon J.-C., Iovanna J.L.;
RT	"Rapid PCR cloning and sequence determination of the rat lithostathine gene.";
RT	Biochim. Biophys. Acta 1174:99-102(1993).
RL	[4]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Miscar;
RA	Miyashita H., Suzuki Y., Watanabe T., Uno M., Morizumi S., Yonekura H., Okamoto H.;
RA	"Structure and characterization of rat Reg I gene.";
RL	Seikagaku 65:1082-1082(1993).
RN	[5]
RP	SEQUENCE OF 22-69.
RP	TISSUE=Pancreas;
RX	MEDLINE=90031455; PubMed=2680252;
RA	Adich Z., de Caro A.M., Guidoni A.A., Mondettra M.B., Rovery M.;
RT	"Characterization in rat pancreatic juice of a protein homologous to the human pancreatic stone protein.";
RL	Comp. Biochem. Physiol. 93B:793-797(1989).
CC	-I- FUNCTION: Might act as an inhibitor of spontaneous calcium carbonate precipitation.
CC	-I- TISSUE SPECIFICITY: Expressed only in regenerating islets, but not in normal pancreatic islets, insulinomas or regenerating liver.
CC	-I- SIMILARITY: Contains 1 C-type lectin family domain.
CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/).
CC	or send an email to license@isb-sib.ch .
CC	-----
DR	EMBL; L07512; AAA41533.1; -
DR	EMBL; M62930; AAA41974.1; -
DR	EMBL; M18962; AAA42028.1; -
DR	EMBL; D26164; BAA05149.1; -
DR	PIR; A28351; A28351.
DR	HSSP; P05451; ILIT.
DR	RCD; 3552; Reg1.
DR	InterPro; IPRO01304; Lectin_C.
DR	InterPro; IPRO03990; Pancreatic_ac.
DR	Pfam; PF00059; Lectin_C_1.
DR	PRINTS; PRO1504; PNCREATITSAF.
DR	SMART; SM00034; CLECT; 1.
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW	Direct protein sequencing; Glycoprotein; Lectin; signal.
FT	SIGNAL 1 21
FT	CHAIN 22 165 Lithostathine.
FT	DOMAIN 33 163 C-type lectin.
FT	DISULFD 35 46 By similarity.
FT	DISULFD 63 161 By similarity.
FT	DISULFD 136 153 By similarity.
FT	CARBOHD 129 129 N-linked (GlcNAc...) (Potential).
SO	SEQUENCE 165 AA; 18672 MW; 9B61EB236B82CF8A CRC64;
Query Match	29.0%; Score 254.5; DB 1; Length 165;
Beet Local Similarity	30.6%; Pred. NO. 5.2e-17;

Matches	52:	Conservative	38:	Mismatches	51:	Indels	29:	Gaps	7:					
QY	7	RLALLSLGLATGYLGDITMRPS-----			CAPGMYTHKSNCGYRKLRLNW				51					
DB	5	KYFILLSTCL-----			MTLSPSQGEAEEDLPGARITCPGSGNAVSVCYFMEDHLNW				56					
QY	52	SDAELECGSYNGAHLSILSLKCASTIAEYI--			SGYQRSPQPIWIGLHDPQKRGQWOWID				109					
DB	57	AEALPFCQNNNSG--YLVSVLSQAQGNFLASIKESGTTAA--			VWIGLHDPQKRRMRHWS				114					
QY	110	GAMTYLRSM--SGKSMGNGN--HCAEMSSNNNFLTSSNECNRRQFLCKYR							157					
DB	115	GSFLYLSKMDYGPNNNSRGTCVSVTNSGYKKRRDNSCDAQLSFVCKFK							164					
RESULT 12														
ID	STR2_STRCA	STANDARD;	PRT;	142 AA.										
AC	P83515;													
DT	05-JUL-2004 (Rel. 44, Created)													
DT	05-JUN-2004 (Rel. 44, Last sequence update)													
DT	05-JUL-2004 (Rel. 44, Last annotation update)													
DS	Struthiolectin-2 (SCA-2).													
DE	Struthio camelus (Ostrich).													
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;													
OC	Archaeosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;													
OX	NCBI_TaxID=8801;													
RN	[1]													
RP	SEQUENCE, SUBCELLULAR LOCATION, PHOSPHORYLATION SITES SER-62; SER-66													
RP	AND SER-68, AND MASS SPECTROMETRY.													
RC	YISSU=Eggshell matrix;													
RC	PubMed=14726203; DOI=10.1016/j.bbapap.2003.09.006;													
RA	Mann K., Siedler F.;													
RT	Ostrich (Struthio camelus) eggshell matrix contains two different C-													
RT	type lectin-like proteins. Isolation, amino acid sequence, and													
RT	posttranslational modifications.";													
RL	Biochim. Biophys. Acta 1696:41-50(2004).													
CC	-1- SUBCELLULAR LOCATION: Eggshell matrix.													
CC	-1- MASS SPECTROMETRY: MW=16834.1; MW_ERR=2; METHOD=Electrospray;													
CC	RANGE=1-142; NOTE=Ref.1.													
CC	-1- SIMILARITY: Contains 1 C-type lectin family domain.													
DR	InterPro; IPR001304; Lectin_C.													
DR	InterPro; IPR003990; Pancreatc.ac.													
DR	Pfam; PF00059; lectin C; 1.													
DR	PRINTS; PRO1504; PNCREATITSAF.													
DR	SMART; SMO0034; CLECT; 1.													
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; PALS_NRG.													
DR	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.													
KW	Direct protein sequencing; Lectin; Phosphorylation.													
FT	DOMAIN 13													
FT	DISULFID 6													

DB 123 ATWDELCSDRKPFICEYR 141

RESULT 13

PAP3_MOUSE STANDARD, PRT, 174 AA.

AC 009049;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DB Pancreatitis-associated protein 3 precursor (REG III-gamma).

GN Name=Pap3; Synonyms=Reg3g;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Liver, and Pancreas;

RX MEDLINE=97208868; PubMed=9055810; DOI=10.1016/S0378-1119(96)00589-6;

RA Narushima Y., Umno M., Nakagawa K.-I., Mori M., Miyashita H.,

RA Suzuki Y., Noguchi N., Takasawa S., Kumagai T., Yonekura H.,

RA Okamoto H.;

RT "Structure, chromosomal localization and expression of mouse genes encoding type III Reg, RegIII alpha, RegIII beta, RegIII gamma.";

RL Gene 185:159-168(1997).

CC -!- FUNCTION: Might be a stress protein involved in the control of bacterial proliferation.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Constitutively expressed in the small intestine, moderately in colon and at an extremely low level in healthy pancreas.

CC -!- SIMILARITY: Contains 1 C-type lectin family domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D63361; BAA18930.1; -.

DR EMBL; D63362; BAA18931.1; -.

DR HSSP; P05451; 11RT.

DR MGP; MGI:109406; Reg3g.

DR InterPro; IPR001304; Lectin C.

DR InterPro; IPR003990; Pancreatis_ac.

DR Pfam; PF00059; Lectin C; 1.

DR PRINTS; PR01504; PNCREATITSAP.

DR SMART; SM00034; CLECT; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

DR Acute phase; Inflammatory response; Lectin; Multigene family; signal.

FT SIGNAL 1 26 Potential.

FT CHAIN 27 174 Pancreatitis-associated protein 3.

FT DOMAIN 38 174 C-type lectin.

FT DISULFID 40 51 By similarity.

FT DISULFID 68 170 By similarity.

FT DISULFID 145 162 By similarity.

FT SEQUENCE 174 AA; 19307 MW; 5575B956A4D8CEP CRC64;

Query Match 28.1%; Score 246.5; DB 1; Length 174;

Best Local Similarity 35.2%; Pred. No. 3.4e-16;

Matches 57; Conservative 24; Mismatches 66; Indels 15; Gaps 6;

QY 10 LILSCL-AKTGVLDII-----MRPSCAPGMFYHKSNCYGYFRKLRNMSDAELRCQSYG 62

DB 13 MLTSLMLSLQYQGVAKKDASSRSRSCPKSRAGSYCYALFVSXKMYDMDMACQCRP 72

QY 63 NGAHLSLILSLKEASTIAEYI-SGYORSQPIWIGLHP-----QKRQOMQWIDGAMLYR 116

DB 73 SG-HLVSVYLSGAELASFLSSMIKSSGNSGQYVWIGLHDPITLGYEPNRGGMWNSADVMNYI 131

QY 117 SW-SGKSGKGNKHCAMSSNNFLTWSSNECKRQHFLCKYR 157

DB 132 NMEINPSSSGNHCCTLSRASGFLKMEYCNLELPYCKRK 173

RESULT 14

O6TR86 PRELIMINARY, PRT, 164 AA.

AC O6TR86;

DT 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE Btcul.

OS Bothrops jararacussu (Jararacussu).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidae;

OC Viperidae; Crotalinae; Bothrops.

OX NCBI_TaxID=8726;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Venom;

RX PubMed=15135412; DOI=10.1016/j.pep.2004.02.012;

RA Kaasab B.H., de Carvalho D.D., Oliveira M.A., Baptista G.R.,

RA Pereira G.A., Novello J.C.;

RT "Cloning, expression, and structural analysis of recombinant Btcul, a c-type lectin from the Bothrops jararacussu snake venom.";

RL Protein Expr. Purif. 35:344-352(2004).

DR EMBL; AY388642; AAQ2957.1; -.

DR GO; GO:0005529; F:sugar binding; IEA.

DR InterPro; IPR001304; Lectin C.

DR InterPro; IPR003990; Pancreatis_ac.

DR Pfam; PF00059; Lectin C; 1.

DR PRINTS; PR01504; PNCREATITSAP.

DR SMART; SM00034; CLECT; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

FT CHAIN 24 157 Btcul.

FT SEQUENCE 164 AA; 19070 MW; F8810D4338B94DC CRC64;

Query Match 27.9%; Score 245; DB 2; Length 164;

Best Local Similarity 36.9%; Pred. No. 4.5e-16;

Matches 48; Conservative 16; Mismatches 62; Indels 4; Gaps 2;

QY 29 SCAPGMFYHKSNCYGYFRKLRNMSDAELRCQSYNGAHLAIIISLKEASTIAEYISGYOR 88

DB 25 NCPQDMLPMLNGLCYKIRNELKAMDAEMFCRKYRPGCHLASIHLYGESPEIAEYISDYHK 84

QY 89 SQ-PIWIGLHPQKRQOMQWIDGAMLYRSWSGKS---MGGNKHCAMSSNNFLTWSSN 144

DB 85 GQSEWVIGLCQKKDFSWEMWDRSCDTYLSMDKNQPHYQKRCVLELVSVTGYRLMNDQ 144

QY 145 ECKRQHFLC 154

DB 145 VCBSKNAFLC 154

RESULT 15

O6OX33 PRELIMINARY, PRT, 158 AA.

AC O6OX33;

DT 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DB C-type lectin.

OS Bothrops insularis (Island jararaca) (Queinada jararaca).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidae;

OC Viperidae; Crotalinae; Bothrops.

OX NCBI_TaxID=8723;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Venom gland;

RA Guimaraes-Gomes V., Oliveira-Carvalho A.L., Pujol-Luz M., Castro H.C.,
RA Junqueira-de-Azevedo I.L.M., Dutra D.L.S.,
RL Ho P.L., Zingali R.B.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522720; AAS01426.1; -
DR GO; GO:0005529; F.sugar binding; IEA.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatic_ac.
DR Pfam; PF00059; Lectin_C; 1.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR01504; PNCREATITISAP.
DR SMART; SM00034; CLECT. 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 158 AA; 18636 MW; 0FA6303E34967EE4 CRC64;

Query Match 27.7%; Score 243.5; DB 2; Length 158;
Best Local Similarity 34.0%; Pred. No. 6e-16;
Matches 51; Conservative 22; Mismatches 68; Indels 9; Gaps 3;
QY 10 ILLSCLAKTGVLDIIMRPSGAPGWFYHKSNCYGFPRKLRNWSDAELQCQSYGNGAHLAS 69
DB 11 LLVVFLSLSGAKGN-----NCPQDWLPNNGLCYKIFDELKAWDAEMPCRYKPGCHLAS 65
QY 70 ILSLKEASTIAEYISGYQRSQ-PIWIGLHDPQRQOWIDGAMVLYRSGWSGKS--MGG 125
DB 66 FHLYGSPETIAEYISDYHKGQSEVWIGLWDKKDFSEWETDRSCTDYLSMDKNQPDHYON 125
QY 126 NHCHEMSNNNPLTWSSNECNKROHFLCK 155
DB 126 KEPCVELVSDTGYRLMNDQVCESKNAFLCQ 155

Search completed: February 11, 2005, 23:32:39
Job time : 179 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 23:32:47 ; Search time 39 Seconds

(without alignments)
389,801 Million cell updates/sec

Title: US-09-525-041-2

Sequence: 1 MASRSMRLLLLSCLAKTGV.....LTWSSNECKMKRQFLCKTRP 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

1: PIR1:1*
2: PIR2:1*
3: PIR3:1*
4: PIR4:1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254.5	29.0	165	A47148	reg I, regenerating
2	254.5	29.0	165	A28351	pancreatic stone p
3	241.5	27.5	142	S78596	ovocleidin - chick
4	240	27.3	166	RGHUIB	regenerating islet
5	236.5	26.9	173	B47148	reg II, regenerati
6	236.5	26.9	174	I83377	regenerating prote
7	233	26.5	135	A38609	lectin, galactose
8	231	26.3	174	S54979	pancreatic islet
9	221	26.0	175	A37194	pancreatic thread
10	228.5	26.0	174	A48689	pancreatic islet
11	225.5	25.7	166	RGHUIA	regenerating islet
12	224	25.5	166	A45751	pancreatic stone p
13	222.5	25.3	172	S32489	lectin - Iberian r
14	221	25.2	175	A49616	pancreatic islet
15	206.5	23.5	152	UC7134	agglutinin alpha
16	203.5	23.2	131	UC5058	bitiscetin alpha c
17	203	23.1	175	A41719	pancreatic stone p
18	202.5	23.1	123	UC2415	echinectin beta cha
19	202.5	23.1	152	UC4690	coagulation factor
20	200.5	22.8	125	UC5059	agglutinin beta cha
21	196.5	22.4	146	UC7105	agglutinin beta cha
22	196.5	22.4	146	UC4691	coagulation factor
23	194	22.1	175	S29822	pancreatic islet
24	193	22.0	330	T46256	brevican - human (
25	193	22.0	312	A54423	brevican precursor
26	191	21.8	883	S57653	brevican precursor
27	189	21.5	133	A47267	botrocetin alpha c
28	185.5	21.1	125	B47267	botrocetin beta ch
29	185.5	21.1	129	UC4329	coagulation factor

30	185	21.1	883	2	S49126	brevican precursor
31	185	21.1	1257	2	S28764	neurocan precursor
32	184.5	21.0	146	2	UC1135	agglutinin beta
33	182.5	20.8	301	2	S13165	asialoglycoprotein
34	182	20.7	144	2	PC7027	agglutinin alpha cha
35	181	20.6	291	1	LNHUI	hepatic lectin H1
36	181	20.6	3562	2	A47171	chondroitin sulfat
37	179	20.4	1268	2	S52781	neurocan - mouse
38	179	20.4	2397	1	A55535	brevican precursor
39	179	20.4	2409	1	A60979	asialoglycoprotein
40	178.5	20.3	311	1	LNHUI2A	asialoglycoprotein
41	177.5	20.2	162	1	LNHUI	lectin BRA3-1 prec
42	177.5	20.2	162	1	LNHUI	lectin BRA3-2 prec
43	177	20.2	1643	2	T14274	verican precursor
44	177	20.2	3381	2	T42389	verican precursor
45	172.5	19.6	123	2	B42972	coagulation factor

ALIGNMENTS

RESULT 1
A47148
reg I, regenerating islet cells - mouse
C.Species: Mus musculus (house mouse)
C.Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 03-Jul-2004
C.Accession: A47148
R.Umo, M.; Yonekura, H.; Nakagawa, K.; Watanabe, T.; Miyashita, H.; Morizumi, S.; Ok
J. Biol. Chem. 268, 15974-15982, 1993
A.Title: Structure, chromosomal localization, and expression of mouse reg genes, reg I a
A.Reference number: A47148; MUID:93340209; PMID:8340418
A.Accession: A47148
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-165 <UNN>
A.Cross-references: UNIPROT:P43137; GB:D14010; NID:G391771; PIDN:BA03111.1; PID:G391772
A.Citations: 21/1, 60/3, 106/3, 144/1
C.Superfamily: tetraenein; C-type lectin homology
P.35-161/Domain: C-type lectin homology <LCH>
P.35-46/63-161, 136-153/Dileptide bonds: #status predicted

Query Match
Best Local Similarity 29.0%; Score 254.5; DB 2; Length 165;
Matches 52; Conservative 35; Mismatches 51; Indels 29; Gaps 7;

QY	10	LLLSCLAKTGVLDIIRP-----SCAPGFYHKSNCYGPRKLRNWSDA	54
DB	8	ILLSCL-----IVLSPQGEAEEDLPASRISCPGSSNAYSCYFTEDRLTWADA	59
QY	55	ELTEGSGYNGAHLASILSKASTIAEYI--SGYRSQPIWIGLHPQKQOQWIDGAM	112
DB	60	DLFCQNMNSG-YLVSVLSQAEQNFVSLIKESGTTDAN-VMTGLDHPKRNRRHWSGSL	117
QY	113	VYRSG-SGKSMGK-NHCAEMSSNNFLTSSNCKNRQHLCKYR	157
DB	118	FLYKSWATGSPNNSNRGYCVSLTSMYTKMKWDNDCAQYSPVCKPK	164

RESULT 2
A28351
pancreatic stone protein precursor - rat
N.Alternate names: lithoecathine
C.Species: Rattus norvegicus (Norway rat)
C.Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C.Accession: A28351; A39081; PLO147; S34618
R.Terazono, K.; Yamamoto, H.; Takasawa, S.; Shiga, K.; Yonemura, Y.; Tochino, Y.; Okamoto
J. Biol. Chem. 263, 2111-2114, 1988
A.Title: A novel gene activated in regenerating islets.
A.Reference number: A92704; MUID:88115343; PMID:2963000
A.Accession: A28351
A.Molecule type: mRNA
A.Residues: 1-165 <TER>

A:Cross-references: UNIPROT:P10758; GB:M18962; NID:g206604; PIDN:AAA42028.1; PID:g206605
R:Rouquier, S.; Verdier, J.M.; Iovanna, J.; Dagorn, J.C.; Giorgi, D.
J. Biol. Chem. 266: 786-791, 1991
A:Title: Rat pancreatic stone protein messenger RNA. Abundant expression in mature exocr
A:Reference number: A39081; MUID:91093273; PMID:1985964
A:Accession: A39081
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-165 <R0U>
A:Cross-references: GB:M2930; GB:J05722; NID:g206462; PIDN:AAA41974.1; PID:g206463
R:Adich, Z.; De Caro, A.M.; Guidoni, A.A.; Moudetir, M.E.; Rovey, M.
Comp. Biochem. Physiol. B 93, 793-797, 1989
A:Title: Characterization in rat pancreatic juice of a protein homologous to the human P
A:Reference number: P10147; MUID:90031455; PMID:2680252
A:Accession: P10147
A:Molecule type: protein
A:Residues: 22-69 <ADR>
A:Experimental source: pancreas
R:Dugetti, N.J.; Frigerio, J.M.; Dagorn, J.C.; Iovanna, J.L.
Biochim. Biophys. Acta 1174, 99-102, 1993
A:Title: Rapid PCR cloning and sequence determination of the rat lithostathine gene.
A:Reference number: S34618; MUID:93326645; PMID:7916640
A:Accession: S34618
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-165 <DUS>
A:Cross-references: EMBL:L07512; NID:g393208; PIDN:AAA41533.1; PID:g393209
C:Comment: This protein is found in pancreatic calculi of mammals. A peptide bond betwe
protein into an insoluble protein at a neutral pH of 5.5 to 7.5.
C:Genetics:
A:introns: 21/1; 60/3; 106/3; 144/1
C:Superfamily: tetraneurin; C-type lectin homology
C:Keywords: pyroglyutamic acid
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-165/Product: pancreatic stone protein #status predicted <MAT>
F:35-161/Domain: C-type lectin homology <LCH>
F:122/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen

Query Match 29.0%; Score 254.5; DB 2; Length 165;
Best Local Similarity 30.6%; Pred. No. 9,2e-18;
Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;
Qy 7 RLALLSLCIATGVLDITMRPS-----CAPGFYHKSNCYFRKLRNW 51
Db 5 KYFILLSTL-----WVLSPSQGEABEDLPARITCPESGNAYSYCYFMEDHLSW 56
Qy 52 SDAELEQSYNGAHLASILSLKEASTIAEYI--SGYRSQPIWGLHPQKQOWMID 109
Db 57 AEADLFCCNMNSG-YLVSVLSQAEQNPLASLKEGTTAAAN-VWIGLHDPKNNRHWSS 114
Qy 110 GAMVLYRSM-SGKSWGK-NHCAEMSSNNNPLTWSSNECNKRQHPFLCKYR 157
Db 115 GSLFLYKSWDTGYPPNNSRNGYCVSTYSNGYKMKMDSCDALSLFVCKFK 164

RESULT 3
S78596
Ovocleidin - chicken
C:Species: Gallus gallus (chicken)
C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 20-Sep-1999
C:Accession: S78596
R:Baernholdt, D.; Andersen, S.O.
submitted to the Protein Sequence Database, September 1998
A:Reference number: S78596
A:Accession: S78596
A:Molecule type: protein
A:Residues: 1-142 <BAR>
A:Experimental source: egg-shell
C:Superfamily: tetraneurin; C-type lectin homology
C:Keywords: phosphoprotein
F:5-16,33-18,113-130/Disulfide bonds: #status experimental
F:61,67/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 27.5%; Score 241.5; DB 2; Length 142;
Best Local Similarity 33.8%; Pred. No. 1,5e-16;
Matches 46; Conservative 23; Mismatches 56; Indels 11; Gaps 3;
Qy 30 CAPGFYHKSNCYFRKLRNWSDAELEQSYNGAHLASILSLKEASTIAEYI----- 83
Db 5 CGPGMVPTPGGLCFPFSRELWSRAESFCRHWGPGSHLAARRSAEIRLLAELTNASRG 64
Qy 84 --SGYRSQPIWGLHPQKQOWIDGAMVLYRSM--SGKSWGK-NHCAEMSSNNNPL 139
Db 65 DGSBEGADGRVWIGLHHPAGSRMSDGTAPRAPSWHRTAKARRGR-CALNDEEAFY 123
Qy 140 TWSSNECNKRQHPFLCK 155
Db 124 SMAARPCTERNAFYCK 139

RESULT 4
RQHUIB
regenerating islet lectin 1-beta precursor - human
N:Alternate names: reg-related protein; regl-beta protein
N:Contents: pancreatic stone protein (PSP)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S34591; S42729; A44712
R:Barthol, C.; Gharib, B.; Giorgi, D.; Sansonetti, A.; Dagorn, J.C.; Berge-Lefranc, J.L.
FEBS Lett. 327, 289-293, 1993
A:Title: A gene homologous to the reg gene is expressed in the human pancreas.
A:Reference number: S34591; MUID:93351647; PMID:8348956
A:Accession: S34591
A:Molecule type: DNA
A:Residues: 1-166 <MOR>
A:Cross-references: UNIPROT:P48304; GB:L08010; NID:g307368; PIDN:AAA18204.1; PID:g487726
A>Note: this gene appears to be expressed in pancreas and liver
R:Motilzum, S.; Watanabe, T.; Umno, M.; Nakagawara, K.; Suzuki, Y.; Miyashita, H.; Yonei
Biochim. Biophys. Acta 1217, 199-202, 1994
A:Title: Isolation, structural determination and expression of a novel reg gene, human r
A:Reference number: S42729; MUID:94153997; PMID:8110835
A:Accession: S42729
A:Molecule type: mRNA
A:Residues: 1-166 <MOR>
A:Cross-references: GB:D16816; NID:g474305; PIDN:BA04091.1; PID:g474306
A:Accession: A44712
A:Molecule type: DNA
A:Residues: 1-166 <MO2>
A:Cross-references: GB:D17291; NID:g474307; PIDN:BA04124.1; PID:g474308
C:Comment: This protein is found in pancreatic calculi of mammals. The tryptic-like cleav
C:Genetics:
A:Gene: GDB:REG1B; REG1
A:Map position: 2p12-2p12
A:Cross-references: GDB:342079
A:introns: 22/1; 61/3; 107/3; 145/1
C:Superfamily: tetraneurin; C-type lectin homology
C:Keywords: glycoprotein; lectin; pancreas; pyroglyutamic acid
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-166/Product: regenerating islet lectin 1beta #status predicted <MAT>
F:34-166/Product: pancreatic stone protein #status predicted <MAT>
F:36-162/Domain: C-type lectin homology <LCH>
F:23/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:127/Binding site: carbonylate (Thr) (covalent) #status predicted
F:33-34/Cleavage site: Arg-11e (trypsin) #status predicted
F:36-47,64-162,137-154/Disulfide bonds: #status predicted

Query Match 27.3%; Score 240; DB 1; Length 166;
Best Local Similarity 34.3%; Pred. No. 2,5e-16;
Matches 46; Conservative 27; Mismatches 57; Indels 4; Gaps 4;
Qy 27 RPSCAPGFYHKSNCYFRKLRNWSDAELEQSYNGAHLASILSLKEASTIAEYISGY 86
Db 33 RISCPBEGNAVRSCYFNEBPERVVDADLYCCNMNSG-NIVSVLTQAEQFVASLIKES 91
Qy 87 QRSQP-IWIGLHDPKQOWMIDGAMVLYRSM-SGKSWGK-NHCAEMSSNNNPLTWSS 143

Db 92 STDSDNVWIGLHDPKRRRHHSSGSLVYKSWDTGSSPSSANAGYCASLTSCSGFKKMD 151

Qy 144 NECNKRQHFLCKR 157
| : : : : :
152 ESCEKFKSFVCKFK 165

RESULT 5

B47148

reg II, regenerating islet cells - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

C:Accession: B47148

R:Ueno, M.; Yonekura, H.; Nakagawara, K.; Watanabe, T.; Miyashita, H.; Morizumi, S.; Ok

J. Biol. Chem. 268, 15974-15982, 1993

A:Title: Structure, chromosomal localization, and expression of mouse reg genes, reg I &

A:Reference number: A47148; MUID:93340209; PMID:8340418

A:Accession: B47148

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-173 <UNN>

C:Cross-references: UNIPROT:Q08731; GB:D14011; NID:9391773; PIDN:BA03112.1; PID:9391774

C:Genetic:

A:introns: 22/1; 68/3; 114/3; 152/1

C:Superfamily: tetranectin; C-type lectin homology

F:43-169/Domain: C-type lectin homology <LCH>

F:43-54,71-169,144-161/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 26.9%; Score 236.5; DB 2; Length 173;

Matches 50; Conservative 33; Mismatches 72; Indels 19; Gaps 5;

Qy 1 MASRSRRLLLLSCLAKTGVLDIT-----MRPSCAPGFYHKNSCYCYFRK 47
| : : : : :
1 MAQNNVYLLFLCLMFLYSQGVABEDPLAEKDLPSAKINCPGANNVGSYCYLLID 60

Db 48 LNWMSDAELSCSYGGAHASTLSLKEASTIAEYI--SGYRSGPIWIGLHDPQKQGM 105
| : : : : :
61 RLTWGADLPCCQM-VAGHIVSLTSGAESNFVNSLVEKSGTASN-VWVGHLDPKSNRRM 118

Qy 106 QWIDGAMLYRSWS--GKSMGNKHCAMSSNNNFLTWSNECNKRQHFLCKR 157
| : : : : :
119 HNSGSLFLFKSNATGAPSTANNGYCVLSLTNAYKMDENCEAGYSVCKFR 172

Db 119 HNSGSLFLFKSNATGAPSTANNGYCVLSLTNAYKMDENCEAGYSVCKFR 172

RESULT 6
183377
regenerating protein III (reg III) - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004

C:Accession: I60296; 183377

R:Suzuki, Y.; Yonekura, H.; Watanabe, T.; Ueno, M.; Morizumi, S.; Miyashita, H.; Okamoto

Gene 144, 315-316, 1994

A:Title: Structure and expression of a novel rat RegIII gene.

A:Reference number: I60296; MUID:94314238; PMID:8039722

A:Accession: I60296

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-174 <RES>

A:Cross-references: UNIPROT:P35231; GB:D23676; NID:9471157; PIDN:BA04904.1; PID:9471158

A:Accession: I83377

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-174 <RES>

A:Cross-references: GB:D26078; NID:9471159; PIDN:BA05071.1; PID:9471160

A:Gene: RegIII

A:introns: 25/1; 64/3; 110/3; 153/1

C:Superfamily: tetranectin; C-type lectin homology

F:39-170/Domain: C-type lectin homology <LCH>

Query Match

Best Local Similarity 26.9%; Score 236.5; DB 2; Length 174;

Matches 35.3%; Pred. No. 5.8e-16;

Matches 49; Conservative 25; Mismatches 56; Indels 9; Gaps 4;

Qy 27 RPSCAPGFYHKNSCYGFRKLRWMSDAELSCSYGGAHASTLSLKEASTIAEYIG- 85
| : : : : :
36 RSCPMGKRAYNSCYTTLVTLKSFQADLCQKRPQ-GLVSLISGSEAFVSSLVYGR 94

Qy 86 YRSQPIWIGLHDPKQKQO-----QWIDGAMLYRSWSG--KSMGNKHCAMSSNNNF 138
| : : : : :
95 VANNQDIIWIGLHDPFMGQDPNGGCHWNSDYVLYNLMDGPSSIVYNRNGCSLRTSEF 154

Db 139 LTWSSNECNKRQHFLCKR 157
| : : : : :
155 LKMGDHDVLELPYCKFR 173

Qy 139 LTWSSNECNKRQHFLCKR 157
| : : : : :
155 LKMGDHDVLELPYCKFR 173

RESULT 7
A38609
lectin, galactose-specific - western diamondback rattlesnake

C:Species: Crotalus atrox (western diamondback rattlesnake)

C>Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 09-Jul-2004

C:Accession: A38609

R:Hiroyoshi, T.; Kasai, K.

J. Biol. Chem. 266, 2320-2326, 1991

A:Title: Complete primary structure of a galactose-specific lectin from the venom of the

A:Reference number: A38609; MUID:91115849; PMID:1989986

A:Accession: A38609

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-135 <HIR>

A:Cross-references: UNIPROT:P21863

C:Superfamily: tetranectin; C-type lectin homology

F:3-111/Domain: C-type lectin homology <LCH>

F:3-11,31-131,106-123/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 26.5%; Score 233; DB 2; Length 135;

Matches 45; Conservative 18; Mismatches 64; Indels 4; Gaps 2;

Qy 29 SCAPGFYHKNSCYGFRKLRWMSDAELSCSYGGAHASTLSLKEASTIAEYISGR 88
| : : : : :
2 NCPDLWLPWNGLCYITFQTKWEDAEFCRYKPKGCHLASFRYGESLEIAEYISDYHK 61

Db 89 SQP-IWIGLHDPKQKQOQWIDGAMLYRSWSGKS--MGKHCAMSSNNNFLTWSN 144
| : : : : :
62 GQENWIGLHDPKQKQPSWETDRSCTDYLTMDKQPDHYNKPFVELVSLGYRLMDQ 121

Qy 145 ECNKRQHFLCK 155
| : : : : :
122 VCESDAFLCQ 132

Db 122 VCESDAFLCQ 132

RESULT 8
S54979
pancreatitis-associated protein PAP-3 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004

C:Accession: S54979; S43438

R:Dusec, N.J.; Frigerio, J.M.; Szpirer, C.; Dagorn, J.C.; Iovanna, J.L.

Biochem. J. 307, 9-16, 1995

A:Title: Cloning, expression and chromosomal localization of the rat pancreatitis-associated

A:Reference number: S54979; MUID:95534061; PMID:7717998

A:Accession: S54979

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-174 <DUS>

A:Cross-references: UNIPROT:P42854; EMBL:U09193; NID:9483931; PIDN:AAA79231.1; PID:94839

R:Frigerio, J.M.; Dusec, N.J.; Garrido, P.; Dagorn, J.C.; Iovanna, J.L.

Biochim. Biophys. Acta 1216, 329-331, 1993

A:Title: The pancreatitis associated protein III (PAP III), a new member of the PAP gene

A:Reference number: S43438; MUID:94060113; PMID:8241280

A:Accession: S43438

A:Status: preliminary

A:Molecule type: mRNA

A:Status: preliminary

A:Molecule type: mRNA

```
A:Residues: 1-174 <FRI>
A:Cross-references: GB:L20869; GB:S67496; NID:g463279; PIDN:AAA41809.1; PID:g463280
C:Genetics:
A:Introns: 26/1; 65/3; 111/3; 153/1
A:Superfamily: tetranectin; C-type lectin homology
F:40-170/Domain: C-type lectin homology <LCH>

Query Match      26.3%; Score 231; DB 2; Length 174;
Best Local Similarity 32.4%; Pred. No. 2e-15;
Matches 55; Conservative 28; Mismatches 73; Indels 14; Gaps 5;

OY 1 MASRSMILLISLACTAGVLGD-----IWRPSCAPQMFYKNSCYGFRKLNMWSDA 54
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 5 VALTTSMWLLISLMLLSQVGEDAKEDVPTSRISCPKGSRAVGSYCYALFVSRSKMFDA 64

OY 55 ELECGSYNGAHLASITLSKEASTIAEYI-SGYQSRQPIWIGLHDPQKROQ-----WQMI 108
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 DLACQKRPBG-HLVSVLGSSEASFPVSLIKSSGNGQVWVIGLHDPQLGQEPNRGMEWS 123

OY 109 DGAMLYRSW-SGKSMGNKHCAMSSNNNPLTWSSNECNKQKHFCLKYR 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 124 NADVNNYNNWETNPSSVSGSHCGTITRASGLRMKNNCISBLPYVCKFK 173

RESULT 9
A37194
pancreatic thread protein precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 09-Jul-2004
C:Accession: A37194; A53897
R:de la Monte, S.M.; Ozuturk, M.; Wands, J.R.
J. Clin. Invest. 86, 1004-1013, 1990
A:Title: Enhanced expression of an exocrine pancreatic protein in Alzheimer's disease at
A:Reference number: A37194; MUID:90368981; PMID:2394826
A:Accession: A37194
A:Molecule type: mRNA
A:Residues: 1-175 <DEU>
A:Cross-references: UNIPROT:P23132; GB:M59794; NID:G163648; PIDN:AAA30750.1; PID:G163648
R:Cal, L.; Harris, W.R.; Marchak, D.R.; Grose, J.; Crabbe, J.W.
J. Protein Chem. 9, 623-632, 1990
A:Title: Structural analysis of bovine pancreatic thread protein.
A:Reference number: A53897; MUID:91197388; PMID:2085387
A:Accession: A53897
A:Molecule type: protein
A:Residues: 38-138/141-175 <CAI>
C:Comment: The purified protein undergoes a reversible globule-fibril transformation and
C:Superfamily: tetranectin; C-type lectin homology
C:Keywords: disulfide bond; extracellular protein; pancreas
F:38-138/Product: pancreatic thread protein chain A #status experimental <ACH>
F:40-171/Domain: C-type lectin homology <LCH>
F:141-175/Product: pancreatic thread protein chain B #status experimental <BCH>

Query Match      26.3%; Score 231; DB 2; Length 175;
Best Local Similarity 30.9%; Pred. No. 2e-15;
Matches 51; Conservative 30; Mismatches 62; Indels 22; Gaps 6;

OY 10 ILLSLACTAGVLGD-----IMRPSCAPQMFYKNSCYGFRKLNMWSDALECQ 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 13 MLSLSLM--LLSQIGENSQKELPSARISCPSSGMAVRSKCYALFKPKYTMMDADICQ 69

OY 60 SYNGAHLASITLSKEASTIAEYISGYQSRQ-PIWIGLHDPQKROQ-----WQMI 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 70 KRPSG-HLVSVLGSSEASFPVSLVNNNLTQGDVIGLHDPFEGSEANAGMEWISNVL 128

OY 114 LYRSW-SGKSMGNKHCAMSSNNNPLTWSSNECNKQKHFCLKY 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 129 NYVAMETDPAISPSSVSGSHCGTITRASGLRMKNNCISBLPYVCKFK 173

RESULT 10
A48689
pancreatic-is-associated protein PAP-2 - rat
C:Species: Rattus norvegicus (Norway rat)
```

```
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A48689
R:Frigerio, J.M.; Duestel, N.J.; Keim, V.; Dgorn, J.C.; Iovanna, J.L.
Biochemistry 32, 9236-9241, 1993
A:Title: Identification of a second rat pancreatic-is-associated protein. Messenger RNA c
A:Reference number: A48689; MUID:93378971; PMID:8369291
A:Accession: A48689
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-174 <FRI>
A:Cross-references: UNIPROT:P35231; GB:L10229; NID:g409014; PIDN:AAA02980.1; PID:g409015
C:Superfamily: tetranectin; C-type lectin homology
F:39-170/Domain: C-type lectin homology <LCH>
F:39-50,67-170,145-162/Disulfide bonds: #status predicted

Query Match      26.0%; Score 228.5; DB 1; Length 174;
Best Local Similarity 34.5%; Pred. No. 3.6e-15;
Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;

OY 27 RPSCAPQMFYKNSCYGFRKLNMWSDALECQSYNGAHLASITLSKEASTIAEYISG- 85
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 36 RTSCPMKAVRSYCYTLVTLKSWFQADLACQKRPBG-HLVSVLGSSEASFPVSVLYGR 94

OY 86 YQSQPIWIGLHDPKROQ-----WQIDGAMLYRSWSG--KSMGNKHCAMSSNNNF 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 95 VNNQDIIWTLHDPQKQPVGGGMEWSNDVLYLWMDGDPSSSTVNRNGCSILTATSEF 154

OY 139 LTVSSNECNKQKHFCLKYR 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 155 LKMDHCHDVBLPFVCKFK 173

RESULT 11
RGHUA
regenerating islet lectin 1-alpha precursor [validated] - human
N:Contains: pancreatic stone protein (PSP)
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1990 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
A:Accession: A35197; B28351; S12950; S02767; S02419; S0013; S01471; A25246
R:Macnabe, T.; Yonekura, H.; Terazono, K.; Yamamoto, H.; Okamoto, H.
J. Biol. Chem. 265, 7432-7439, 1990
A:Title: Complete nucleotide sequence of human reg gene and its expression in normal and
A:Reference number: A35197; MUID:90237042; PMID:232435
A:Accession: A35197
A:Molecule type: DNA
A:Residues: 1-166 <WAT>
A:Cross-references: UNIPROT:P05451; GB:J05412
R:Terazono, K.; Yamamoto, H.; Takasawa, S.; Shiga, K.; Yonemura, Y.; Tochino, Y.; Okamoto,
J. Biol. Chem. 263, 2111-2114, 1988
A:Title: A novel gene activated in regenerating islets.
A:Reference number: A92704; MUID:86115343; PMID:2965000
A:Accession: B28351
A:Molecule type: mRNA
A:Residues: 1-166 <TER>
A:Cross-references: GB:M18963; NID:g190978; PIDN:AAA36558.1; PID:g190979
R:Itch, T.; Tsuzuki, H.; Katch, T.; Teraoka, H.; Matsumoto, K.; Yoshida, N.; Terazono, K
FEBS Lett. 272, 85-88, 1990
A:Title: Isolation and characterization of human reg protein produced in Saccharomyces c.
A:Reference number: S12950; MUID:91032149; PMID:2226837
A:Accession: S12950
A:Molecule type: protein
A:Residues: 23-52/160-166 <ITO>
A>Note: sequence determined from protein isolated after human cDNA sequence was cloned at
R:de Caro, A.M.; Adrich, Z.; Fournet, B.; Capon, C.; Bonicel, J.J.; de Caro, J.D.; Rover
Biochim. Biophys. Acta 994, 281-284, 1989
A:Title: N-terminal sequence extension in the glycosylated forms of human pancreatic sto
A:Reference number: S02767; MUID:89150292; PMID:2493268
A:Accession: S02767
A:Molecule type: protein
A:Residues: 23-47 <DBC>
R:Rouimi, P.; de Caro, J.; Bonicel, J.; Rover, M.; de Caro, A.
```


A:Accession: A45751
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-166 <GIO>
A:Cross-references: UNIPROT:P05451; GB:M27J90; NID:g623412; PIDN:AAA60546.1; PID:g623413
C:/Superfamily: tetranectin; C-type lectin homology
F:/36-162/Domain: C-type lectin homology <LCH>

Query Match 25.5%; Score 224; DB 2; Length 166;
Best Local Similarity 33.3%; Pred. No. 9,4e-15;
Matches 45; Conservative 26; Mismatches 58; Indels 6; Gaps 4;

Dy 27 RPSCAPMFHKNSKCYGFRKLNMWDALKECSYNGAHLASILSKASTAEVY--S S 84
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 33 RISCEBTANRYACVFYNEDRETVDADLYCONMMSG-NLVSVLTQABEAFASLIKES 91

Dy 85 GYGRSPWITGLHDPOKRQQOWIDGAMYLRSW--SGKMGKNHCAEMSSNNFLTWS 142
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 92 GTDDFN-VWIGLDHPKKNRHMHWSSGSLSVGKSMGTGAPSSVPDGYCVSLTSSTGFQMK 150

Dy 143 SNECNKRQHPLCKYR 157
Db 151 DVPCDEKFSFCVKFK 165

RESULT 13

S32489
lectin - Iberian ribbed newt
C:Species: Pleurodeles waltlil (Iberian ribbed newt)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
A:Accession: S32489; S28530
R:Tiffoche, C.; Chennel, A.; Jégo, P.; le Pennec, J.P.
A>Title: Isolation and characterization of a cDNA clone encoding a pleurodeles lectin.
A:Reference number: S32489; PMID:93279340; PMID:8504829
A:Accession: S32489
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-172 <TLF>
A:Cross-references: UNIPROT:Q02988; EMBL:X69062; NID:g64257; PIDN:CAA48800.1; PID:g64258
C:/Superfamily: tetranectin; C-type lectin homology
F:/37-164/Domain: C-type lectin homology <LCH>

Query Match 25.3%; Score 222.5; DB 2; Length 172;
Best Local Similarity 34.3%; Pred. No. 1,4e-14;
Matches 46; Conservative 22; Mismatches 55; Indels 11; Gaps 4;

Dy 30 CAPGMFHKSNCYGFRKLNMWDALKECSYNGAHLASILSKASTAEVYISGVQR 89
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 37 CTPEMDCHFNYSYKXLYPNASWTDAEFYCQLYPGAHLAISHEDENDPLTEITPKNSN 96

Dy 90 QP-IWIGLHDPKRQQOWIDGAMYLY---RSWSGKMGKNHCAEMSSNNFLT--WS 142
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 97 YPVVMWGSSGYCDRSRVWTDGSDMDYQKMRQMPEPSTGREGFCIDE----NFVTPELMN 152

Dy 143 SNECNKRQHPLCKY 156
Db 153 DEHCDCRFPICKY 166

RESULT 14

A49616
pancreatitis-associated protein precursor - human
N:Alternate names: C-type lectin; pancreatic stone protein homolog HIP
C:Species: Homo sapiens (man)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
A:Accession: A49616; A49931; S29821; S48197; I55580
R:Dusetti, N.O.; Fritgerio, J.M.; Fox, M.F.; Swallow, D.M.; Dagorn, J.C.; Iovanna, J.L.
A>Title: Molecular cloning, genomic organization, and chromosomal localization of the hu
A:Reference number: A49616; PMID:94245143; PMID:8188210
A:Molecule type: DNA

A:Residues: 1-175 <DUS>
A:Cross-references: UNIPROT:006141; GB:J15533; NID:9482908; PIDN:AAA60020.1; PID:9482908
R:LaSserre, C.; Chetisa, L.; Simon, M.T.; Vernier, P.; Brehot, C.
Cancer Res. 52, 5089-5095, 1992
A:Title: A novel gene (HIP) activated in human primary liver cancer.
A:Reference number: A44931; MUID:92386513; PMID:1325291
A:Accession: A44931
A:Molecule type: mRNA
A:Residues: 1-175 <LAS>
A:Cross-references: GB:X68641; NID:9312806; PIDN:CAA48605.1; PID:9312807
A:Experimental source: hepatocellular carcinoma
A>Note: Sequence extracted from NCBI backbone (NCBIN:113007, NCBIPI:113008)
R:Itch, T.; Teraoka, H.
Biochim. Biophys. Acta 1172, 184-186, 1993
A:Title: Cloning and tissue-specific expression of cDNA for the human and mouse homolog
A:Reference number: S29821; MUID:93176807; PMID:7679928
A:Accession: S29821
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-175 <LTO>
A:Cross-references: GB:DJ3510; NID:9285970; PIDN:BA02728.1; PID:9285971
R:LaSserre, C.; Simon, M.T.; Ishikawa, H.; Diriong, S.; Nguyen, V.C.; Christa, L.; Vernier, J. Biochem. 224, 29-38, 1994
A:Title: Structural organization and chromosomal localization of a human gene (HIP/PAP)
A:Reference number: S48197; MUID:94357229; PMID:8076648
A:Accession: S48197
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-175 <LA2>
R:Orelle, B.; Kelm, V.; Masciotra, L.; Dagorn, J.C.; Iovanna, J.L.
J. Clin. Invest. 90, 2284-2291, 1992
A:Title: Human pancreaticitis-associated protein. Messenger RNA cloning and expression in
A:Reference number: I55580; MUID:93107309; PMID:1469087
A:Accession: I55580
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-175 <RES>
A:Cross-references: GB:S51768; NID:9262368; PIDN:AB24642.1; PID:9262369
C:Genetics:
A:Gene: GDB:PAP; HIP
A:Cross-references: GDB:136839; OMIM:167805
A:Map position: 2p12-2p12
A:Introns: 26/1; 65/3; 111/3; 154/1
C:Superfamily: tetraneurin; C-type lectin homology
C:Keywords: acute phase; extracellular protein; pancreas
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-175/Product: pancreaticitis-associated protein #status predicted <MAT>
F:40-171/Domain: C-type lectin homology <LCH>
F:40-51,68-171,146-163/Disulfide bonds: #status predicted

Query Match 25.2%; Score 221; DB 2; Length 175;
Best Local Similarity 32.7%; Pred. No. 2e-14; Mismatches 64; Indels 16; Gaps 7;
Matches 53; Conservative 29; Mismatches 64; Indels 16; Gaps 7;

OY 10 LLLSCLAKTVGLGDIIMR--PS---CAPGMFYHKSNCYGFPRKLRNMSDALEECOSYSG 62
Db 13 MLTSLCMLLSQVQGEPEPRLPSARIRCPKSKAYGSHCVLFLSPKSWTDADLACQKRP 72

OY 63 NGALHASTLSTKEASTIAFYISGYORQSP--IWIGLHDPQKROQ---WQWIDGAMVLYR 116
Db 73 SG-MIVSVLISGAEPSFVSLKSGNSYSYVWIGLHDPQGTGTPNGEGEWSSSDVMNYF 131

OY 117 SM--SGKSMGNKKAEMSSNNNFTWSSNECNKRQHFLCKY 156
Db 132 AWRNPSTISSPGHCASLSRSTAFLRWDVNCVNLPTVCKF 173

RESULT 15
JC7134
agkiscutacin alpha chain precursor - sharp-nosed viper
N:Alternate names: fibrinogenolytic venom protein
C:Species: Agkistrodon acutus (sharp-nosed viper)
C:Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 09-Jul-2004

C:Accession: JC7134; PC7037
R:Cheng, X.; Qian, Y.; Liu, Q.; Li, B.X.Y.; Zhang, M.; Liu, J.
Biochem. Biophys. Res. Commun. 265, 530-535, 1999
A:Title: Purification, characterization, and cDNA cloning of a new fibrinogenolytic venom
A:Reference number: JC7134; MUID:20025379; PMID:10558903
A:Accession: JC7134
A:Molecule type: mRNA
A:Residues: 1-152 <CHE>
A:Cross-references: UNIPROT:Q9DEF9; UNIPROT:Q91AM1; UNIPROT:Q8J1W0; GB:AF176420
A:Experimental source: venom gland
A:Accession: PC7037
A:Molecule type: protein
A:Residues: 24-53; 84-86; 87-94; 125-136; 137-152 <CH2>
C:Superfamily: tetraneurin; C-type lectin homology
C:Keywords: disulfide bond; heterodimer; venom
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-152/Product: agkiscutacin alpha chain #status experimental <MAT>

Query Match 23.5%; Score 206.5; DB 2; Length 152;
Best Local Similarity 31.3%; Pred. No. 4.5e-13;
Matches 47; Conservative 26; Mismatches 64; Indels 13; Gaps 4;

OY 10 LLLSCLAKTVGLGDIIMRPSCAPGMFYHKSNCYGFPRKLRNMSDALEECOSYNGAHLAS 69
Db 11 LTVVFLSLSGTAAD-----CSSGSSSYEGHCYKVPKQSKTWADAESEFCTQVNGCHLVS 64

OY 70 ILSLKEASTIAFYISGYORQSP--IWIGLHDPQKROQ---WQWIDGAMVLYRSGKSMNG 125
Db 65 IESSGENDFVAHLIAQKIKSAKIHWIGLRAQNEKQCSLEWSSSIYENWIEB--- 121

OY 126 NKHCAEMSSNNNFTWSSNECNKRQHFLCK 155
Db 122 SKKGLGHIETGFIKMFENFYCEQDDPFVCE 151

Search completed: February 11, 2005, 23:43:48
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 23:21:46 ; Search time 49 Seconds
(without alignments)
1053.599 Million cell updates/sec

Title: US-09-525-041-2

Perfect score: 878
Sequence: 1 MASRMRLLILSLCAKTGV.....LTWSNECKRQHPICKRP 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13:	/cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
17:	/cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
18:	/cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep:*
19:	/cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
20:	/cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878	100.0	158	US-09-922-217-1070	Sequence 1070, Ap
2	878	100.0	158	US-09-922-217-1077	Sequence 1077, Ap
3	878	100.0	158	US-09-922-217-1078	Sequence 1078, Ap
4	878	100.0	158	US-09-922-217-1079	Sequence 1079, Ap
5	878	100.0	158	US-09-922-217-1080	Sequence 1080, Ap
6	878	100.0	158	US-09-833-263-1070	Sequence 1070, Ap
7	878	100.0	158	US-09-833-263-1077	Sequence 1077, Ap
8	878	100.0	158	US-09-833-263-1078	Sequence 1078, Ap
9	878	100.0	158	US-09-833-263-1079	Sequence 1079, Ap
10	878	100.0	158	US-09-833-263-1080	Sequence 1080, Ap
11	878	100.0	158	US-09-525-041-2	Sequence 2, Appl
12	878	100.0	158	US-09-969-034-4471	Sequence 4471, Ap
13	878	100.0	158	US-10-025-380-1070	Sequence 1070, Ap

14	878	100.0	158	US-10-025-380-1077	Sequence 1077, Ap
15	878	100.0	158	US-10-025-380-1078	Sequence 1078, Ap
16	878	100.0	158	US-10-025-380-1079	Sequence 1079, Ap
17	878	100.0	158	US-10-025-380-1080	Sequence 1080, Ap
18	878	100.0	158	US-10-100-6089-2	Sequence 2, Appl1
19	878	100.0	158	US-10-157-027-114	Sequence 114, Ap
20	878	100.0	158	US-10-205-823-347	Sequence 347, Ap
21	878	100.0	158	US-10-295-027-138	Sequence 138, Ap
22	878	100.0	158	US-10-295-027-181	Sequence 781, Ap
23	878	100.0	158	US-10-295-027-861	Sequence 861, Ap
24	878	100.0	158	US-10-295-027-1200	Sequence 1200, Ap
25	878	100.0	158	US-10-173-999-933	Sequence 93, Appl
26	878	100.0	158	US-10-734-564-105	Sequence 105, Ap
27	878	100.0	158	US-10-099-7815-2	Sequence 2, Appl1
28	878	100.0	166	US-10-106-639-6394	Sequence 6394, Ap
29	760	86.6	167	US-10-452-646-9	Sequence 9, Appl1
30	240	27.3	174	US-09-925-297-753	Sequence 753, Ap
31	231	26.3	174	US-10-028-248A-113	Sequence 113, Ap
32	231	26.3	174	US-10-107-782-113	Sequence 113, Ap
33	231	26.3	175	US-10-028-248A-110	Sequence 110, Ap
34	231	26.3	175	US-10-107-782-110	Sequence 110, Ap
35	225.5	25.7	166	US-09-997-003-43	Sequence 43, Appl
36	225.5	25.7	166	US-09-997-003-43	Sequence 43, Appl
37	225.5	25.7	166	US-10-734-564-4	Sequence 4, Appl1
38	225.5	25.7	174	US-09-925-301-1182	Sequence 1182, Ap
39	224	25.5	166	US-10-734-564-2	Sequence 2, Appl1
40	223	25.4	175	US-10-028-248A-113	Sequence 38, Appl
41	223	25.4	175	US-10-107-782-113	Sequence 38, Appl
42	221	25.2	175	US-10-316-761-3	Sequence 3, Appl1
43	221	25.2	175	US-10-484-906-1	Sequence 1, Appl1
44	221	25.2	175	US-10-028-248A-109	Sequence 109, Ap
45	221	25.2	175	US-10-107-782-109	Sequence 109, Ap

ALIGNMENTS

RESULT 1
US-09-922-217-1070
; Sequence 1070, Application US/09922217
; Patent No. US-00200764141
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongrong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1070
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1070

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MASRMRLLILSLCAKTGVLDIMRPSCAPGMFYHNSNCYGFRKLRMSDALEQCS 60
|||||

Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRLKLRMSDAELBCQS 60
QY 61 YGNGAHLASISLKEASTIAEYISGYORSQPIWIGLHDPQRQOWIDGAMTYLRWSWG 120
Db 61 YGNGAHLASISLKEASTIAEYISGYORSQPIWIGLHDPQRQOWIDGAMTYLRWSWG 120
QY 121 KSMGNKHCAMSSNNNFLTWSNNECNKRQHFLLCKYRP 158
Db 121 KSMGNKHCAMSSNNNFLTWSNNECNKRQHFLLCKYRP 158

RESULT 2

US-09-922-217-1077
; Sequence 1077, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1077
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1077

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRLKLRMSDAELBCQS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRLKLRMSDAELBCQS 60
QY 61 YGNGAHLASISLKEASTIAEYISGYORSQPIWIGLHDPQRQOWIDGAMTYLRWSWG 120
Db 61 YGNGAHLASISLKEASTIAEYISGYORSQPIWIGLHDPQRQOWIDGAMTYLRWSWG 120
QY 121 KSMGNKHCAMSSNNNFLTWSNNECNKRQHFLLCKYRP 158
Db 121 KSMGNKHCAMSSNNNFLTWSNNECNKRQHFLLCKYRP 158

RESULT 3

US-09-922-217-1078
; Sequence 1078, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.

; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1078
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1078

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRLKLRMSDAELBCQS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRLKLRMSDAELBCQS 60
QY 61 YGNGAHLASISLKEASTIAEYISGYORSQPIWIGLHDPQRQOWIDGAMTYLRWSWG 120
Db 61 YGNGAHLASISLKEASTIAEYISGYORSQPIWIGLHDPQRQOWIDGAMTYLRWSWG 120
QY 121 KSMGNKHCAMSSNNNFLTWSNNECNKRQHFLLCKYRP 158
Db 121 KSMGNKHCAMSSNNNFLTWSNNECNKRQHFLLCKYRP 158

RESULT 4

US-09-922-217-1079
; Sequence 1079, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1079
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1079

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRLKLRMSDAELBCQS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRLKLRMSDAELBCQS 60
QY 61 YGNGAHLASISLKEASTIAEYISGYORSQPIWIGLHDPQRQOWIDGAMTYLRWSWG 120
Db 61 YGNGAHLASISLKEASTIAEYISGYORSQPIWIGLHDPQRQOWIDGAMTYLRWSWG 120

QY 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158
Db 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 5

US-09-922-217-1080
; Sequence 1080, Application US/09922217
; Patent No. US20020764141
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yudi
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1080
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1080

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGMFYHKSNCYGYFRKLRNMSDALEECOS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGMFYHKSNCYGYFRKLRNMSDALEECOS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPTWIGLHDPQRQOWIDGAMTYLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPTWIGLHDPQRQOWIDGAMTYLYRSWSG 120
QY 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158
Db 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 6

US-09-833-263-1070
; Sequence 1070, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1070
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-833-263-1070

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGMFYHKSNCYGYFRKLRNMSDALEECOS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGMFYHKSNCYGYFRKLRNMSDALEECOS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPTWIGLHDPQRQOWIDGAMTYLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPTWIGLHDPQRQOWIDGAMTYLYRSWSG 120
QY 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158
Db 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 7

US-09-833-263-1077
; Sequence 1077, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1077
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1077

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGMFYHKSNCYGYFRKLRNMSDALEECOS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGMFYHKSNCYGYFRKLRNMSDALEECOS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPTWIGLHDPQRQOWIDGAMTYLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPTWIGLHDPQRQOWIDGAMTYLYRSWSG 120
QY 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158
Db 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 8

US-09-833-263-1078
; Sequence 1078, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1078
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1078

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Qy 61 YGNGAHLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKRQOWIDGAMLYLRWSWG 120
Db 61 YGNGAHLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKRQOWIDGAMLYLRWSWG 120
Qy 121 KSMGNGHCAEMSSNNNFLLTWSSNECNKRQHFLLCKYRP 158
Db 121 KSMGNGHCAEMSSNNNFLLTWSSNECNKRQHFLLCKYRP 158

RESULT 9
US-09-833-263-1079
; Sequence 1079, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1079
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1079

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Qy 61 YGNGAHLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKRQOWIDGAMLYLRWSWG 120
Db 61 YGNGAHLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKRQOWIDGAMLYLRWSWG 120
Qy 121 KSMGNGHCAEMSSNNNFLLTWSSNECNKRQHFLLCKYRP 158
Db 121 KSMGNGHCAEMSSNNNFLLTWSSNECNKRQHFLLCKYRP 158

RESULT 10
US-09-833-263-1080
; Sequence 1080, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1080
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1080

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Qy 61 YGNGAHLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKRQOWIDGAMLYLRWSWG 120
Db 61 YGNGAHLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKRQOWIDGAMLYLRWSWG 120
Qy 121 KSMGNGHCAEMSSNNNFLLTWSSNECNKRQHFLLCKYRP 158
Db 121 KSMGNGHCAEMSSNNNFLLTWSSNECNKRQHFLLCKYRP 158

RESULT 11
US-09-525-041-2
; Sequence 2, Application US/09525041
; Publication No. US20030158098A1
; GENERAL INFORMATION:
; APPLICANT: Soppet et al.
; TITLE OF INVENTION: Colon Specific Gene and Protein
; FILE REFERENCE: P178D2
; CURRENT APPLICATION NUMBER: US/09/525,041
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 09/162,508
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: US 08/468,413
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: patentin version 3.0
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-525-041-2

Query Match 100.0%; Score 878; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Qy 61 YGNGAHLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKRQOWIDGAMLYLRWSWG 120
Db 61 YGNGAHLASIIISLKEASTIAEYISGYORSOPITWIGLHPQKRQOWIDGAMLYLRWSWG 120
Qy 121 KSMGNGHCAEMSSNNNFLLTWSSNECNKRQHFLLCKYRP 158
Db 121 KSMGNGHCAEMSSNNNFLLTWSSNECNKRQHFLLCKYRP 158

RESULT 12
US-09-969-034-4471
; Sequence 4471, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:

```

; APPLICANT: Burgess, Christopher C.
; APPLICANT: Aetle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthachi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4471
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-034-4471

Query Match          100.0%; Score 878; DB 11; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGRKLRNMSDALEECOS 60
DB 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGRKLRNMSDALEECOS 60

QY 61 YNGGAHLASITSLKASTIAEYISGYSQPSQPIWIGLHDPQKQOQOWIDGAMYLYRSWSG 120
DB 61 YNGGAHLASITSLKASTIAEYISGYSQPSQPIWIGLHDPQKQOQOWIDGAMYLYRSWSG 120

QY 121 KSMGNKHCACEMSSNNNFLTWSNNECNKRQHFLLCKYRP 158
DB 121 KSMGNKHCACEMSSNNNFLTWSNNECNKRQHFLLCKYRP 158

RESULT 13
US-10-025-380-1070
; Sequence 1070, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yudi
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yaeli A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1070
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1070
```

```

Query Match          100.0%; Score 878; DB 13; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGRKLRNMSDALEECOS 60
DB 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGRKLRNMSDALEECOS 60

QY 61 YNGGAHLASITSLKASTIAEYISGYSQPSQPIWIGLHDPQKQOQOWIDGAMYLYRSWSG 120
DB 61 YNGGAHLASITSLKASTIAEYISGYSQPSQPIWIGLHDPQKQOQOWIDGAMYLYRSWSG 120

QY 121 KSMGNKHCACEMSSNNNFLTWSNNECNKRQHFLLCKYRP 158
DB 121 KSMGNKHCACEMSSNNNFLTWSNNECNKRQHFLLCKYRP 158

RESULT 14
US-10-025-380-1077
; Sequence 1077, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yudi
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yaeli A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1077
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1077

Query Match          100.0%; Score 878; DB 13; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGRKLRNMSDALEECOS 60
DB 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGRKLRNMSDALEECOS 60

QY 61 YNGGAHLASITSLKASTIAEYISGYSQPSQPIWIGLHDPQKQOQOWIDGAMYLYRSWSG 120
DB 61 YNGGAHLASITSLKASTIAEYISGYSQPSQPIWIGLHDPQKQOQOWIDGAMYLYRSWSG 120

QY 121 KSMGNKHCACEMSSNNNFLTWSNNECNKRQHFLLCKYRP 158
DB 121 KSMGNKHCACEMSSNNNFLTWSNNECNKRQHFLLCKYRP 158

RESULT 15
US-10-025-380-1078
; Sequence 1078, Application US/10025380
; Publication No. US20020182191A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1078
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1078

Query Match      100.0%; Score 878; DB 13; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MASRSMRLLLSLSCAKTGVGDIIMRSPCAPGWFYHKSNCYGYFRKLRNMSDALEECQS 60
Db      1 MASRSMRLLLSLSCAKTGVGDIIMRSPCAPGWFYHKSNCYGYFRKLRNMSDALEECQS 60

QY      61 YGNGNHLASIIISLKEASTIAEYISGYGRSOPWIGLHDPOKROQWQWIDGAMTYLYRSNSG 120
Db      61 YGNGNHLASIIISLKEASTIAEYISGYGRSOPWIGLHDPOKROQWQWIDGAMTYLYRSNSG 120

QY      121 KSMGNGKCAEMSSNNNFLTWSNNECNTRQHFLECKYRP 158
Db      121 KSMGNGKCAEMSSNNNFLTWSNNECNTRQHFLECKYRP 158
```

Search completed: February 11, 2005, 23:33:33
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 23:29:46 ; Search time 43 Seconds

(without alignments)
274.292 Million cell updates/sec

Title: US-09-525-041-2

Sequence: 1 MASRMRLLLSLCAKTGV.....LTWSSNECKRQHFICKRP 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878	100.0	158	2	US-08-729-103-1
2	878	100.0	158	2	US-08-468-413-2
3	878	100.0	158	3	US-09-162-508-2
4	878	100.0	158	5	PCT-US95-07169-2
5	695	79.2	122	4	US-09-621-976-4812
6	254.5	29.0	165	2	US-08-401-530A-7
7	254.5	29.0	165	2	US-08-729-103-3
8	254.5	29.0	165	2	US-08-709-662-7
9	240	27.3	166	2	US-08-729-103-4
10	240	27.3	174	4	US-09-949-016-10686
11	231	26.3	174	4	US-08-401-530A-5
12	231	26.3	174	2	US-08-709-662-5
13	228.5	26.0	174	2	US-08-401-530A-6
14	228.5	26.0	174	2	US-08-709-662-6
15	224	25.5	144	4	US-09-949-016-10685
16	224	25.5	166	4	US-09-949-016-6286
17	223	25.4	117	6	5514582-15
18	223	25.4	117	6	5514582-15
19	221	25.2	175	2	US-08-464-637-2
20	221	25.2	175	2	US-08-401-530A-4
21	221	25.2	175	2	US-08-709-662-4
22	221	25.2	175	2	US-08-822-261-3
23	221	25.2	175	4	US-09-426-852-3
24	221	25.2	174	4	US-09-949-016-10537
25	218	24.8	174	1	US-07-778-156-7
26	218	24.8	174	2	US-08-822-261-4
27	218	24.8	174	2	US-08-422-166-7

28	218	24.8	174	4	US-09-226-852-4	Sequence 4, Appl1
29	208	23.7	175	2	US-08-401-530A-3	Sequence 3, Appl1
30	208	23.7	175	2	US-08-709-662-3	Sequence 3, Appl1
31	204.5	23.3	130	1	US-07-893-929A-7	Sequence 7, Appl1
32	204.5	23.3	130	5	PCT-US92-10344-7	Sequence 7, Appl1
33	204	23.2	175	2	US-08-822-261-1	Sequence 1, Appl1
34	204	23.2	175	4	US-09-226-852-1	Sequence 1, Appl1
35	198.5	22.6	132	1	US-07-893-929A-5	Sequence 5, Appl1
36	198.5	22.6	132	5	PCT-US92-10344-5	Sequence 5, Appl1
37	198.5	22.6	134	1	US-07-893-929A-2	Sequence 2, Appl1
38	198.5	22.6	134	5	PCT-US92-10344-2	Sequence 2, Appl1
39	196	22.3	125	1	US-07-893-929A-3	Sequence 3, Appl1
40	196	22.3	125	5	PCT-US92-10344-3	Sequence 3, Appl1
41	194.5	22.2	131	1	US-07-893-929A-1	Sequence 1, Appl1
42	194.5	22.2	131	5	PCT-US92-10344-1	Sequence 1, Appl1
43	193	22.0	912	5	PCT-US95-03747-2	Sequence 2, Appl1
44	190	21.6	174	2	US-08-401-530A-2	Sequence 2, Appl1
45	190	21.6	174	2	US-08-709-662-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-729-103-1

Sequence 1 Application US/08729103

Patent No. 583,784

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/729,103

FILING DATE: Filed Herewith

Prior APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36, 749

REFERENCE/DOCKET NUMBER: PF-0138 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 158 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: COMFET02

CLONE: 1310334

US-08-729-103-1

Query Match 100.0%; Score 878; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 5.9e-88;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07169
FILING DATE: 06 JUN 95
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-3699
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1100
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US95-07169-2

```

Query Match	100.0%;	Score 878;	DB 5;	Length 158;
Best Local Similarity	100.0%;	Pred. No. 5.9e-88;		
Matches 158;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

Qy	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200																																																																																
Db	121	KSMGNGKHCHEMSSNNNF.LTWSNSNECNKRQHF.LCKTRP	122		123		124		125		126		127		128		129		130		131		132		133		134		135		136		137		138		139		140		141		142		143		144		145		146		147		148		149		150		151		152		153		154		155		156		157		158		159		160		161		162		163		164		165		166		167		168		169		170		171		172		173		174		175		176		177		178		179		180		181		182		183		184		185		186		187		188		189		190		191		192		193		194		195		196		197		198		199		200	

```

RESULT 5
US-09-621-976-4812
: Sequence 4812 Application US/09621976
: Patent No. 6639063
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J.Y.
: TITLE OR INVENTION: ESTs and Encoded Human Proteins
: FILE REFERENCE: GENSET.054R2
: CURRENT APPLICATION NUMBER: US/09/621,976
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 19335
: SOFTWARE: Patent.pm
: SEQ ID NO 4812
: LENGTH: 122
: TYPE: PRF
: ORGANISM: Homo sapiens
: US-09-621-976-4812

```

Query Match	79.2%;	Score 695;	DB 4;	Length 122;
Best Local Similarity	100.0%;	Pred. No. 3.9e-68;		
Matches 122;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

[illegible]

RESULT 6
 US-08-401-530A-7
 Sequence 7, Application US/08401530A
 Patent No. 5834590
 GENERAL INFORMATION:
 APPLICANT: Vink, Aaron I.
 APPLICANT: Piltenger, Gary L.
 APPLICANT: Rafeaeloff, Ronit
 APPLICANT: Rosenberg, Lawrence
 APPLICANT: Duguid, William P.
 TITLE OF INVENTION: INGP PROTEIN INVOLVED IN PROLIFERATION
 TITLE OF INVENTION: ISLET NEOGENESIS
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Banner & Allegretti
 STREET: 1001 G Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: US
 ZIP: 20001-4597
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/401,530A
 FILING DATE: 22-FEB-1995
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Kagan, Sarah A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 00570,48743
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100
 TELEFAX: 202-508-9239
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 165 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Rattus rattus
 US-08-401-530A-7

Query Match	29.0%;	Score 254.5;	DB 2;	Length 165;
Best Local Similarity	30.6%;	Pred. No. 6.6e-20;		
Matches	52;	Conservative 38;	Mismatches 51;	Indels 29;
				Gaps 7

```

Oy      7 RLILLLSLCAKGVGLADIMRPS-----CAGMWYHKSNCYGRKLRNW 51
      Oy      :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      5 KYIILSL-----MYLSPSGQGEADEDLPSARITCPGGSNAASYCYFFHDLISW 56
      Oy      :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Oy      52 SDAELECGSYGNGAHILASTILKEASTAETI--SGYORSQPIWIGLHDPKROOWID 109
      Oy      :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      57 AEADLPQGNMNSG-YLVSVLSQAEGNPLASTILKESGTTAAN-VWIGLHDPGNRRMHWSS 114
      Oy      :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Oy      110 GAWLYTRSW-SGKSMGGRK-HCAEMSSNNNPLTMSNECNKGRHFLCYKR 157
      Oy      :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```

Db 115 GSLFLYKSMDTGYPNNSNRGYCVSVTSNSGYKKWRDNCDAQLSFVCKFK 164

RESULT 7

US-08-729-103-3

Sequence 3, Application US/08729103

Patent No. 5837841

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Gold, Surya K.

TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FaetsEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/729.103

FILING DATE: Filed Herewith

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0138 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 165 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 393209

US-08-729-103-3

Query Match 29.0%; Score 254.5; DB 2; Length 165;

Best Local Similarity 30.6%; Pred. No. 6.6e-20;

Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;

Qy 7 RLILLSCLAKTVGLGDIIMRPS-----CAPGMFYHNSNCYGFRKLRNW 51

Db 5 KYFILLSC-----MVLSPSGQAEEDLPARITCPBGSNAVSICYTFMEDHLNW 56

Qy 52 SDAELFCQSYGNGAHLASILSLKEASTIAEYI--SGYRSQPIWIGLHDPQKROQWQWID 109

Db 57 AEAULPCQNNMSG-YLVSVLSQAEGNFLASLIKESGTTAAN-VWIGLHDPKNNRRHWSS 114

Qy 110 GAMVLYRSW-SGKSMGKN-HCAEMSSNNNFLTWSNECNKRQHFLLCKYR 157

Db 115 GSLFLYKSMDTGYPNNSNRGYCVSVTSNSGYKKWRDNCDAQLSFVCKFK 164

RESULT 8

US-08-709-662-7

Sequence 7, Application US/08709662

Patent No. 5840531

GENERAL INFORMATION:

APPLICANT: Vink, Aaron I.

APPLICANT: Pittenger, Gary L.

APPLICANT: Rafaeloff, Ronit

APPLICANT: Rosenberg, Lawrence

APPLICANT: Duguid, William P.

TITLE OF INVENTION: INCAP PROTEIN INVOLVED IN PANCREATIC

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.

STREET: 1001 G Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: US

ZIP: 20001-4597

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/709,662

FILING DATE: 09-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 00570.59178

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9239

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 165 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Rattus rattus

US-08-709-662-7

Query Match 29.0%; Score 254.5; DB 2; Length 165;

Best Local Similarity 30.6%; Pred. No. 6.6e-20;

Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;

Qy 7 RLILLSCLAKTVGLGDIIMRPS-----CAPGMFYHNSNCYGFRKLRNW 51

Db 5 KYFILLSC-----MVLSPSGQAEEDLPARITCPBGSNAVSICYTFMEDHLNW 56

Qy 52 SDAELFCQSYGNGAHLASILSLKEASTIAEYI--SGYRSQPIWIGLHDPQKROQWQWID 109

Db 57 AEAULPCQNNMSG-YLVSVLSQAEGNFLASLIKESGTTAAN-VWIGLHDPKNNRRHWSS 114

Qy 110 GAMVLYRSW-SGKSMGKN-HCAEMSSNNNFLTWSNECNKRQHFLLCKYR 157

Db 115 GSLFLYKSMDTGYPNNSNRGYCVSVTSNSGYKKWRDNCDAQLSFVCKFK 164

RESULT 9

US-08-729-103-4

Sequence 4, Application US/08729103

Patent No. 5837841

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Gold, Surya K.

TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

QY 109 DGAMLYRSM-SGKSMGKHKCAEMSSNNNPLTWSSNECNKRQHLCKYR 157
Db 124 NADVMNYFWETNPSSVSGSHCGTLTTRASGFLRMENNCISLPLVYCKFK 173

RESULT 12

US-08-709-662-5
Sequence 5, Application US/08709662
Patent No. 5840531
GENERAL INFORMATION:
APPLICANT: Vink, Aaron I.
APPLICANT: Piltenger, Gary L.
APPLICANT: Rafaeloff, Ronit
APPLICANT: Rosenberg, Lawrence
APPLICANT: Duguid, William P.
TITLE OF INVENTION: INGP PROTEIN INVOLVED IN PANCREATIC
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,662
FILING DATE: 09-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 00570,59178
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
US-08-709-662-5

Query Match 26.3%; Score 231; DB 2; Length 174;

Best Local Similarity 32.4%; Pred. No. 2,6e-17;
Matches 55; Conservative 28; Mismatches 73; Indels 14; Gaps 5;

QY 1 MASRSMRLILSLCAKTVGLD-----IIRPSCAPGMFYHKNCYGFRLRMMSDA 54
Db 5 VALTMSMWLLSLMLLSQVGEDEAKEDVPTSRICSPKRSRAYSCTALPFSVSGSWFDA 64
QY 55 BLECGSYNGAHLASILSKEASTIAEYI-SGYQSRQPIWIGLHDPQKRQ-----WQMT 108
Db 65 DLACQKRRPSG-HLVSVLSSGSBASFVSLIKSSGSGQVWVIGLHDPQLGQEPNRGMEWS 123
QY 109 DGAMLYRSM-SGKSMGKHKCAEMSSNNNPLTWSSNECNKRQHLCKYR 157
Db 124 NADVMNYFWETNPSSVSGSHCGTLTTRASGFLRMENNCISLPLVYCKFK 173

RESULT 13

US-08-401-530A-6
Sequence 6, Application US/08401530A
Patent No. 5834590
GENERAL INFORMATION:

APPLICANT: Vink, Aaron I.
APPLICANT: Piltenger, Gary L.
APPLICANT: Rafaeloff, Ronit
APPLICANT: Rosenberg, Lawrence
APPLICANT: Duguid, William P.
TITLE OF INVENTION: INGP PROTEIN INVOLVED IN PANCREATIC
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401,530A
FILING DATE: 22-FEB-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 00570,48743
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
US-08-401-530A-6

Query Match 26.0%; Score 228.5; DB 2; Length 174;

Best Local Similarity 34.5%; Pred. No. 4.9e-17;
Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;

QY 27 RPSCAPGMFYHKNCYGFRLRMMSDABLECGSYNGAHLASILSKEASTIAEYISG- 85
Db 36 RTSQPMGSKAYRSCYTLVTTLKSWFQADLACQKRRPSG-HLVSVLSSGSASFVSLVYGR 94
QY 86 YQSRQPIWIGLHDPQKRQ-----WQMTDGAMLYRSMG--KSMGKHKCAEMSSNNNF 138
Db 95 VNNQDILWIMLHDTMQQPRGGGMEWSNDVLYNLMMDDGPSSTVNRKNGCSLTATSEF 154
QY 139 LTWSSNECNKRQHLCKYR 157
Db 155 LKMGDHHCDVELPVCCKFK 173

RESULT 14

US-08-709-662-6
Sequence 6, Application US/08709662
Patent No. 5840531
GENERAL INFORMATION:
APPLICANT: Vink, Aaron I.
APPLICANT: Piltenger, Gary L.
APPLICANT: Rafaeloff, Ronit
APPLICANT: Rosenberg, Lawrence
APPLICANT: Duguid, William P.
TITLE OF INVENTION: INGP PROTEIN INVOLVED IN PANCREATIC
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.

STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,662
FILING DATE: 09-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 00570.59178
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
US-08-709-662-6

Query Match 26.0%; Score 228.5; DB 2; Length 174;
Best Local Similarity 34.5%; Pred. No. 4.9e-17;
Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;

QY 27 RPSCAPGWFYHKSNCYGRKLRNMSDAELCQSYNGAHLASILSKASTIAEYISG- 85
DB 36 RISCPEGNAYRSYCYTTLTKSPQADLACQKRPSSG-HLVSLTSGEASPVSLVTR 94
QY 86 YGRSQPIWIGLHDPQKROQ-----WQMDGAMLYRSWSG--KSMGKNGKCAEMSSNNP 138
DB 95 VNNNDIWIWIMHDPMTGQPGNGGEMSSNDVNIYLMNDGDPSSYVNRGCGSLTATSE 154
QY 139 LTWSSNECKRQHFLCKYR 157
DB 155 LKWGDHHCDELPEVCCKF 173

RESULT 15
US-09-949-016-10685
Sequence 10685, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10685
LENGTH: 144
TYPE: PRT
ORGANISM: Human
US-09-949-016-10685

Query Match 25.5%; Score 224; DB 4; Length 144;
Best Local Similarity 33.3%; Pred. No. 1.2e-16;
Matches 45; Conservative 26; Mismatches 58; Indels 6; Gaps 4;

QY 27 RPSCAPGWFYHKSNCYGRKLRNMSDAELCQSYNGAHLASILSKASTIAEYI--S 84
DB 11 RISCPEGNAYRSYCYTTLTKSPQADLACQKRPSSG-HLVSLTSGEASPVSLVTR 69
QY 85 YGRSQPIWIGLHDPQKROQWQMDGAMLYRSW--KSMGKNGKCAEMSSNNPFLWS 142
DB 70 GTDDFN-VWIGLHDPKRRMHTSSGSLVSKWIGIGAPSSVNPQCVSLTSTGFQKWK 128
QY 143 SNECKRQHFLCKYR 157
DB 129 DVPCEDKPSFVCKFK 143

Search completed: February 11, 2005, 23:43:03
Job time : 44 secs

This Page Blank (uspto)